IOWA STATE UNIVERSITY
Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

1989

The molecular analysis of c1-p and c1-m1, two recessive alleles of C1: a regulatory locus of the anthocyanin pathway in the aleurone tissue of Zea mays L.

Brian Eric Scheffler
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

 Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, and the Genetics Commons

Recommended Citation

Scheffler, Brian Eric, "The molecular analysis of c1-p and c1-m1, two recessive alleles of C1: a regulatory locus of the anthocyanin pathway in the aleurone tissue of Zea mays L. " (1989). *Retrospective Theses and Dissertations*. 9176.
https://lib.dr.iastate.edu/rtd/9176

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9014951

The molecular analysis of *c1-p* and *c1-m1*, two recessive alleles of *C1*: A regulatory locus of the anthocyanin pathway in the aleurone tissue of *Zea mays* L.

Scheffler, Brian Eric, Ph.D.

Iowa State University, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

The molecular analysis of *c1-p* and *c1-m1*, two recessive alleles of *C1*: A regulatory locus of the anthocyanin pathway in the aleurone tissue of *Zea mays* L.

by

Brian Eric Scheffler

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Agronomy
Major: Plant Breeding and Cytogenetics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1989

TABLE OF CONTENTS

1. INTRODUCTION

Anthocyanin biosynthesis is a well defined plant biological system. Although several plant species have been used to study anthocyanin genes and their expression, *Zea mays* L. represents one of the best organisms. In maize, there are at least 10 genes known to be involved in anthocyanin biosynthesis in aleurone tissue. Some of the loci code for structural components while others regulate the structural genes (Reddy and Coe, 1962; Styles and Ceska, 1977; McCormick, 1978; Dooner, 1983).

The *C1* locus is an important regulatory locus of the anthocyanin pathway in maize. Although other regulatory loci have been cloned, such as *R1* (Dellaporta et al., 1988) and *Vp1* (McCarty et al., 1989), *C1* is best defined at the molecular and genetic level. *C1* regulates the anthocyanin pathway and has tissue specific expression (expressed only in the aleurone and scutellum tissue of maize kernels). The *C1* protein sequence has homology to *myb* proto-oncogene products, and the protein might act as a transcriptional activator (Paz-Ares et al., 1987).

The structure of *C1*, with its relation to *myb* proto-oncogenes and transcriptional activators, indicates that it may regulate anthocyanin expression at the level of transcription (Paz-Ares et al., 1987, 1989). In addition

to this homology, the C1 locus is important because it has several unique alleles, such as C1-I, c1-p, and c1-m1. The molecular analysis of C1 and its alleles has and will increase our understanding of how regulatory genes function.

The C1-I allele is dominant to all other C1 alleles and inhibits anthocyanin production (color formation). Molecular analysis of C1-I has elicited a probable reason for its dominant inhibitor nature. In this case, it appears that the deletion of the activator domain (acidic domain) turns the protein into a transcriptional suppressor instead of an activator (Paz-Ares et al., 1989).

The study presented here involves the analysis of two recessive C1 alleles (c1-p and c1-m1). In homozygous c1-p mature kernels color is not produced, but those kernels do produce visible anthocyanins when germinated in the presence of light.

The c1-m1 allele was the first mutable allele isolated by McClintock (1948, 1949, 1951). The c1-m1 allele arose from the transposition of the "standard" Ds (which causes chromosome breakage) from its original position into the C1 locus (McClintock, 1948). When this Ds element excises, gene function can either be restored or

chromosome breaks can occur. Preliminary molecular analysis of c1-m1 indicates the Ds element is inserted 2 kb downstream of the translation stop site (Cone et al., 1986; Förger, 1988).

The molecular investigation of these c1-p and c1-m1 may reveal important regions that influence the expression of the C1 locus. Examination of the c1-p allele may reveal an area that interacts with light induced factors, while the analysis of c1-m1 could uncover the importance of the 3' flanking region of C1. It is possible that the Ds transposable element in c1-m1 is inserted in a regulatory region of the C1 locus or that it effects gene expression by preventing the formation of normal RNA transcripts. In addition, investigation of this Ds element should increase our understanding of how Ds elements can cause chromosome breaks.

2. LITERATURE REVIEW

The production of anthocyanins in various plants and their different tissues is important not only because they add color to our lives but they also represent a well defined biological system. The study of the expression of anthocyanins has lead to significant molecular and genetic discoveries in plants (Peterson, 1988).

During the genetic and biochemical analysis of anthocyanin genes in Zea mays several important concepts including paramutation, tissue specific expression and coordinate gene expression in plants were developed. Molecular genetics has taken advantage of the knowledge gained from the genetic and biochemical work in order to isolate transposable elements and to gain a molecular understanding of how gene interactions work.

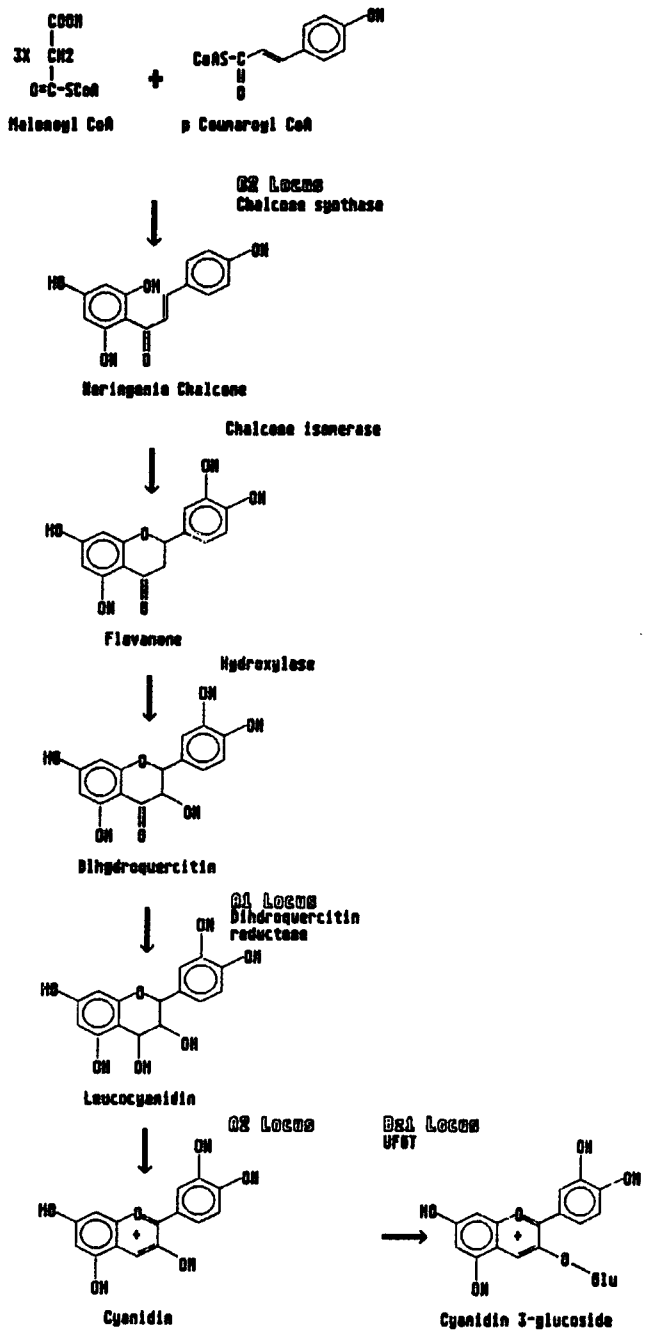
Although several plant species have been used to study anthocyanin genes and their expression (Peterson, 1988), Zea mays represents one of the best analyzed systems. In maize, there are at least 10 genes known to be involved in anthocyanin biosynthesis in aleurone tissue.

2.1. Coordinate Gene Regulation of the Anthocyanin Pathway

Production of anthocyanins in the maize aleurone requires the expression of several genes and a mutation in nearly any of these genes results in the failure of the aleurone to produce anthocyanins (Reddy and Coe, 1962, Styles and Ceska, 1977; McCormick, 1978). Some loci code for structural components in the pathway (A1, A2, Bz1, Bz2, C2 and Pr1), see Figure 2.1, whereas others regulate expression of the structural genes (C1, Dek1, R1, and Vp1) (Dooner, 1983).

Dooner (1983) demonstrated that the regulatory loci previously listed, controlled anthocyanin biosynthesis in a coordinated manner. This was done by measuring the levels of two independent enzymes involved in anthocyanin biosynthesis, chalcone synthase (CHS) and UDPglucose 3-O-glucosyltransferase (UFGT). For each of the anthocyanin loci the measurements were taken on aleurone tissue, which is triploid, from kernels with different gene doses of the wild type allele (+ +/+, + +/-, - -/+, and - -/-), (Dooner and Nelson, 1977 and 1979; Dooner 1983). Kernels that were homozygous recessive for c1, dek1, r1, and vp1 had abnormally low levels of both enzymes, while even one dose of the wild type allele restored enzyme expression to near normal levels. Aleurone tissue with recessive a1, a2, or

Figure 2.1. A diagram of the anthocyanin pathway in the aleurone tissue of *Zea mays*


Malonoyl CoA
p Coumaroyl CoA
C2 Locus
Chalcone synthase
Naringenin Chalcone
Chalcone isomerase
Flavanone
Hydroxylase
Dihydroquercitin
A1 Locus
Dihydroquercitin
reductase
Leucocyanidin
C2 Locus
Cyanidin
Bz1 Locus
UFGT
Cyanidin 3-glucoside
Glu

bz2 alleles always had normal levels of UFGT and CHS, independent of gene dosage. Homozygous bz1 kernels had no UFGT expression and normal levels of CHS. The level of UFGT rose in a linear fashion as the gene dosage of Bz1 increased. C2 kernels had the opposite biochemical phenotype as bz1 in that UFGT levels were always normal and CHS expression responded to an increase in gene dosage.

The data gained from these experiments indicated that C2 and Bz1 code for the enzymes CHS and UFGT, respectively. It was also concluded that C1, Dek1, R1 and Vp1 coordinately controlled C2 and Bz1, because mutations at any of these loci inhibited CHS and UFGT expression (Dooner, 1983). Although these are known as regulatory loci, the mechanism of how they govern anthocyanin biosynthesis is not known.

2.2. Tissue Specific Expression and Duplicated Genes

Genetic and molecular analyses have demonstrated that several of the maize anthocyanin genes are expressed tissue specifically, and that often genes at different loci encode the same function. In this case, different loci are usually, but not exclusively, preferentially expressed in different tissues. Functional genes that encode for the same protein product will be referred to as

duplicated genes in this text. These genes usually have different promoters, thus allowing for tissue specific expression, and their protein products may have a slightly different amino acid constitution although their function is the same.

Of the structural genes, C2 and Whp (white pollen) represent an excellent example of duplicated genes with different tissue specific expression. Both of the loci code for the enzyme chalcone synthase (Figure 2.1) and have the same exon intron structure, but the sequence of the introns and promoters are dissimilar (Niesbach-Klösgen, 1987).

The Whp gene is normally expressed only in plant tissue and not in the kernel. Whp is also expressed in the aleurone tissue when a kernel is homozygous for in (Coe, 1982; P. Franken, Max-Planck-Institut für Züchtungsforschung, personal communication). In this situation, the kernels are either darkly pigmented (presence of C2) or lightly pigmented (homozygous c2). In contrast, C2 seems to be expressed in most tissues. Homozygous recessive c2 plants and kernels do not produce anthocyanins except for specific tissues and under special conditions. Plants that are homozygous for C2 or whp are

fertile, but the double homozygotes produce only sterile white pollen (Coe et al., 1981).

It is clear that C2 and Whp are regulated by different mechanisms and, in the case of Whp, one of the genes regulating it appears to be In (P. Franken, Max-Planck-Institut für Züchtungsforschung, personal communication).

The situation with duplicated regulatory genes is similar to the structural genes, except that the exact nature of how the product functions is usually not known. Pl and C1 are duplicated loci regulating the anthocyanin pathway and their coding regions are about 90% homologous (K. Cone, University of Missouri, personal communication). There seems to be little or no homology in the regions flanking the coding sequence. C1 regulates anthocyanin biosynthesis only in the aleurone and scutellum tissues, while Pl regulates this pathway in most of the other plant tissues (Coe and Neuffer, 1977). Since the coding regions of both genes are similar, it is logical to conclude that they regulate anthocyanin biosynthesis via the same mechanism, and it is only the regulation of their own expression that determines their tissue specific nature.

2.3. The C1 Locus

To understand gene regulation it is important to have a well defined biological system. The anthocyanin pathway

fits this requirement because the biochemistry has been extensively analyzed and it is known that several genes are required for pigment formation. Of all the genes involved, the regulatory loci may be the most important to study since it is possible that they directly regulate the expression of structural loci. Of the regulatory loci, the C1 locus probably represents the best gene for analysis. Unlike R1 which seems to be a very complex locus (Coe and Neuffer, 1977), and Vp1 and Dek1 which are pleiotropic (Dooner, 1983), the C1 locus is relatively simple, but contains some interesting features.

The C1 locus is expressed in a tissue specific manner (aleurone and scutellum), affects only anthocyanin production, and has several interesting mutant alleles. One allele (C1-I) is a color inhibitor and is dominant to all other known C1 alleles. Kernels that have two or three doses of C1-I (C1-I C1-I/C1 or C1-I C1-I/C1-I) are colorless and those with one copy (C1 C1/C1-I) have light pigmentation. Another important mutant is c1-p which is light sensitive. Homozygous c1-p kernels are colorless at maturation, but when germinated in the light the aleurone and scutellum tissues produce anthocyanins (Hsu, 1970; Kirby and Styles, 1970; Chen and Coe, 1977).

As with most of the other anthocyanin loci, there are several transposable element induced mutations at the *C1* locus (Peterson, 1978), which were essential for cloning the locus and beneficial in elucidating the gene's structure (Paz-Ares et al., 1986; Cone et al., 1986). One particular mutant (*c1-m1*) is of historical relevance because it is the first transposable element induced mutable allele isolated by McClintock (1948, 1949, 1951). In addition, *c1-m1* is important in that the transposable element *Ds* is inserted in the 3' region, about 2 kb downstream of the poly(A) addition site (Paz-Ares et al., 1987; Förger, 1988). It is possible this *Ds* element blocks normal gene expression by interrupting RNA transcription, or by destroying a regulatory region (Paz-Ares et al., 1987).

2.3.1. *The c1-p allele*

The mutant allele *c1-p* was defined by Hsu (1970) when she tested different recessive *c1* alleles to see if they were light inducible. She found that some recessive alleles did not respond to light (*c1-negative*, *c1-n*), while others did (*c1-positive*, *c1-p*). Red light of about 650 nm was the most effective wavelength and infrared light inhibited color formation. When *c1-p* plants were crossed to *C1-I* plants, the progeny did not produce color

when germinated in the light indicating that C1-I is dominant to c1-p.

Chen and Coe (1977) conducted a more extensive analysis of c1-p kernels and discovered several interesting genetic and physiological details. Genetically, they were able to show that c1-p was not affected by loci regulating anthocyanin in plant tissue, and that kernels homozygous for c1-p and vp1 (c1-p/c1-p, vp1/vp1) produced anthocyanin when exposed to light. They also showed that anthocyanin production occurs in the aleurone and scutellum tissues.

It is also known that anthocyanin production in kernels homozygous recessive only for vp1 is light sensitive (McCormick, 1977c; Scheffler, 1986). This indicates that light may stimulate a shunt that can bypass C1 and Vp1. c1-p kernels are only light sensitive in the early stages of germination. The photoperiod and the stage of germination when light is applied determine the amount of anthocyanins produced. Chen and Coe (1977) determined (Table 2.1) that if kernels were germinated for four hours, and then exposed to light for an additional four hours, that these conditions were sufficient to stimulate the maximum amount of anthocyanin production. Kernels are significantly less sensitive after 31 hours

Table 2.1. The effect of illumination and germination on c1-p homozygous kernels (Chen and Coe, 1977)

Light duration	Germination time preceding illumination					
	4 hr.	12 hr.	20 hr.	31 hr.	37 hr.	44 hr.
1 hour	7[a]	7	6	2	2	1
4 hours	8	8	7	3	2	1
Beginning to end	8	8	7	5	2	1

[a]The number given is a relative rating of anthocyanin pigmentation, with a range from 0 (no pigmentation) to 8 (intense pigmentation).

and almost insensitive by 48 hours. It was concluded that after 48 hours the aleurone cells may not be physiologically active or they may have a specific period of light sensitivity.

Perhaps one of the most interesting aspects of c1-p is that the kernels can "store" the light induction signal. Maturing c1-p ears can be exposed to light by removing the husk. The mature kernels are colorless, however slight pigmentation will occur if the light source is intense (McCormick, 1977b). The colorless kernels can be germinated in the dark and anthocyanins will accumulate in the aleurone and scutellum tissues (Chen and Coe, 1977).

2.3.2. The Ds induced mutant c1-m1

Maize transposable elements can be divided into two categories, those that can code for their own transposition (regulatory elements) and those that depend on a signal produced by the regulatory element to transpose (receptor elements). Basically, receptor elements can only respond to one type of regulatory element and not another. When this response exists, the receptor and regulatory elements are said to compose a transposable element system.

The first transposable element, Ds of the Ac/Ds transposable element system, was described by McClintock

(1946). In the presence of the regulatory element Ac, the original Ds isolate ("standard Ds") exhibited chromosome breakage at a site just proximal to Wx on the short arm of chromosome 9 (McClintock, 1947, 1948, 1951). The chromosome breaks resulted in the loss of the distal segment. Crosses between C1 sh bz wx +Ac females and C1-I Sh Bz1 Wx Ds males resulted in breaks occurring at the Ds location. The ensuing loss of the acentric fragment exposed the C1 sh bz1 wx phenotype. Therefore, the aleurone layer appeared as a colorless background with bronze sectors (McClintock, 1947, 1948, 1951). In several instances, it seems that the Ds element moved from its standard position to another location on the short arm of chromosome 9 and caused chromosome breaks at its new location (McClintock, 1950). On one occasion, the mutant c1-m1 was isolated (McClintock, 1948, 1949, 1951).

Plants homozygous for c1 sh Bz1 wx (female) and C1 Sh Bz1 wx Ds +Ac (male) were crossed, and an exceptional kernel was found that had colored spots on a colorless background instead of the expected colorless spots on a colored background. The new mutant was designated c1-m1. Genetic analysis demonstrated that c1-m1 was caused by the simultaneous excision of the standard Ds and its insertion in or near the C1 locus of the male parental chromosome.

McClintock (1948, 1949, 1951) demonstrated that in the presence of Ac, the Ds element could excise out of the c1-m1 allele and the C1 phenotype would be restored, or the excision event could cause the chromosome to break.

Chromosome breakage was observed through cytological analysis. McClintock (1949, 1951) proved that the Ds element which caused the c1-m1 mutation also caused the chromosome breakage by examining somatic revertant sectors and germinal revertant kernels (c1-m1 -> C1). That the Ds at c1-m1 was the basis for chromosome breakage could be proven by functional analysis. If a Ds element located in another section of the short arm of chromosome 9 caused the chromosome breaks cytologically observed, then the c1-m1 germinal revertant kernels would still have chromosome breakage. In addition, one would expect chromosome breaks within the revertant somatic sectors, seen as colorless sectors within colored spots on a colorless background. McClintock (1949, 1951) observed that chromosome breaks were no longer detected when the Ds element excised out of the c1-m1 allele.

The combined genetic and cytogenetic evidence correlating the simultaneous loss of the standard Ds with occurrence of the mutant c1-m1, which is conditioned by a Ds insertion that causes chromosome breaks, indicated that

the standard *Ds* element is inserted into *c1-m1* (McClintock, 1949, 1951).

The molecular isolation and evaluation of *Ds* elements indicate that there are three basic classes of elements (Figure 2.2) (Döring and Starlinger, 1984). Class I consists of *Ds* elements that are deletion derivatives of *Ac* elements. Class II are elements that only contain the terminal repeat of the *Ac* element and the internal DNA has no homology to *Ac* (Sutton et al., 1984). Class III are *Double Ds* elements where a Class I *Ds* element is inserted within another Class I *Ds* (Döring and Starlinger, 1984). *Ac* elements have inverted terminal repeats of 11 bp (TAGGGATGAAA and TGGCACCCTG) and all *Ds* elements isolated have the terminal G nucleotide replaced with an A (TGGCACCCTA).

Döring and Starlinger (1984) have proposed that the *Ds* induced chromosome breaks are caused by Double *Ds* elements. The *Double Ds* elements have two pairs of the 11 bp inverted repeats and therefore, direct repeats of the sequence are found adjacent to each other. If the *Ac* transposase cleaves at the direct repeats, chromosome breakage could occur (Figures 2.3 and 2.4).

Figure 2.2. A comparison of the molecular structure of the regulatory element Ac and several Ds receptor elements (modified from Döring and Starlinger, 1984). The symbol "(" indicates the coordinates of a deletion with respect to Ac and the delta sign designates the deletion. The broken and wavy lines indicate sequences that have no homology to Ac or each other. In the Double Ds, the arrows designate the orientation of the Ds elements. In this example, a Ds element is inserted in the opposite orientation to Ds5933 at position 3614, as indicated with an arrow

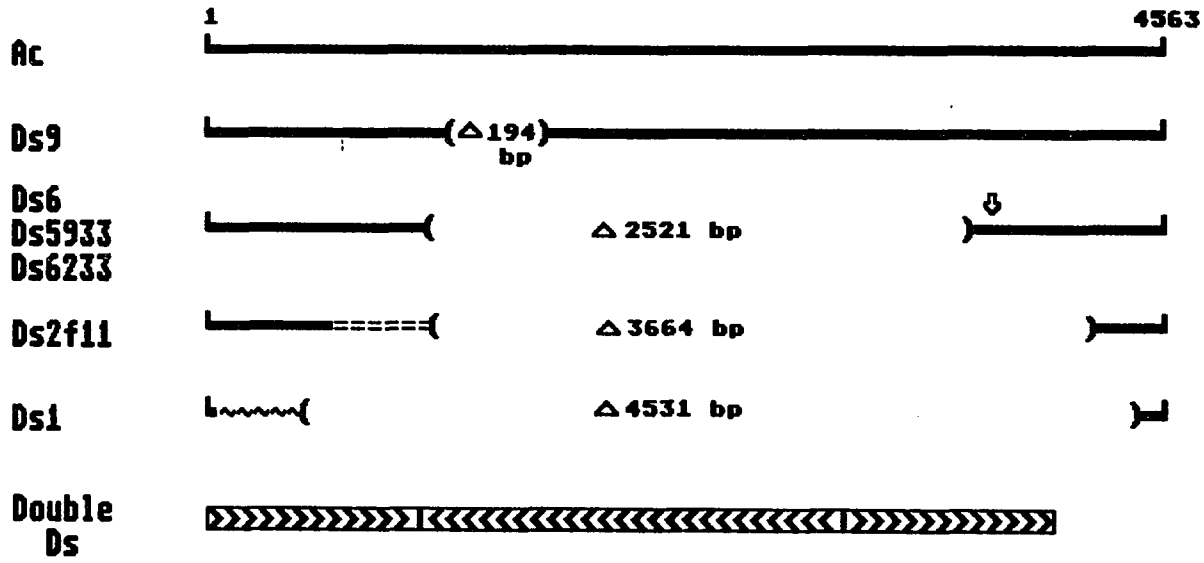

1
4563
Ac
Ds9
(△194)
bp
Ds6
Ds5933
Ds6233
△2521 bp
Ds2f11
△3664 bp
Ds1
△4531 bp
Double
Ds

Figure 2.3. A diagram of the excision of a normal Ds element (modified after Döring and Starlinger, 1984). The black boxes indicate the 8 bp host direct duplication. The open arrows indicate the element's 11 bp inverted repeats. Upon excision, the inverted repeats are aligned and the DNA is cleaved at the site indicated with the black arrows. The Ds element is removed and the chromosomal fragments 1-2 and 3-4 are ligated together

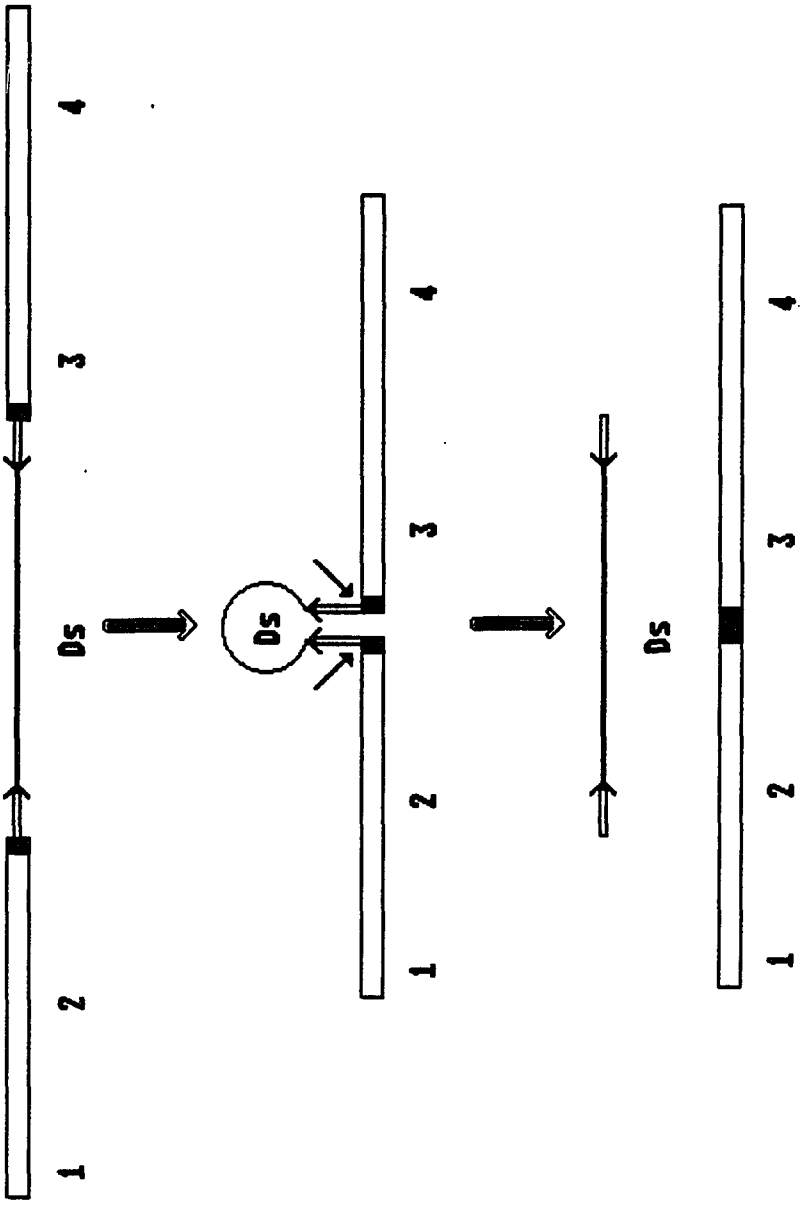
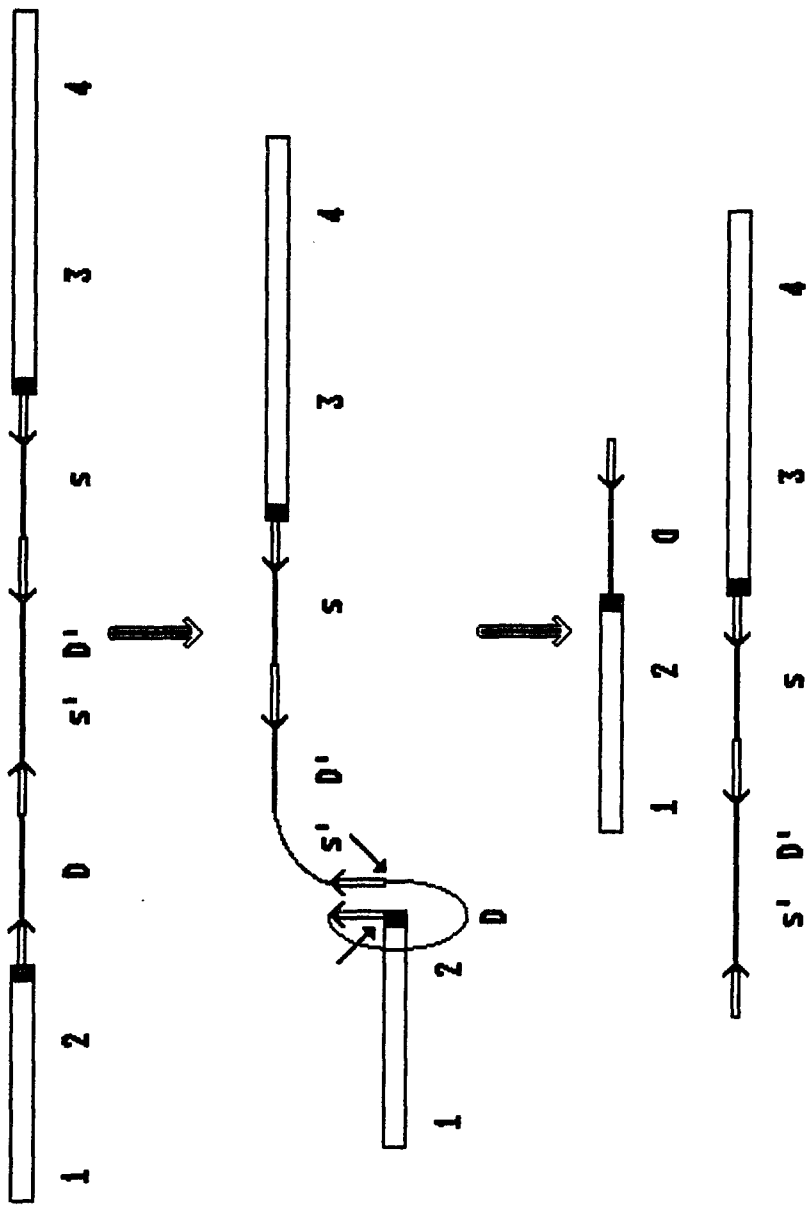

4
3
Ds
2
1
Ds
4
3
2
1
Ds
4
3
2
1

Figure 2.4. A diagram of the excision of a Double Ds element resulting in a chromosome break (modified after Döring and Starlinger, 1984). A second Ds element, indicated as "s' D'", is inserted in an opposite orientation into the first Ds element. The black boxes indicate the 8 bp host direct duplication. The open arrows indicate the element's 11 bp inverted repeats. Upon excision, one inverted repeat from each Ds element is aligned, and the DNA is cleaved at the site indicated with the black arrows. The one end of Ds (D) is ligated to the chromosomal fragment 1-2. Unlike the example in Figure 2.3, the chromosomal segments are not rejoined thus resulting in a chromosome break


1
2
D
S' D'
S
3
4
1
2
D
S' D'
S
3
4
1
2
D
S' D'
S
3
4

2.3.3. Cloning of the wild type C1 allele

One method for cloning a gene is to isolate its protein product. The protein can be sequenced and synthetic oligonucleotides are used as probes to screen a cDNA library. An alternative is raise antibodies against the protein and to use these to screen a cDNA expression library. The major limitations to this procedure are the identification and isolation of the protein. Unfortunately, some loci, and regulatory genes in particular, have low expression and usually no parameters are available by which the protein can be identified.

A popular alternative method is the use of transposon tagging (Bingham et al., 1981). This procedure requires that a defined transposable element is inserted in the desired locus and that it has been previously cloned. By using the transposable element as a probe in a genomic library, the tagged locus can be isolated. This procedure has several limiting factors. For example, mutants with known transposable element inserts must first be identified, and transposable elements can be present at high copy number in the genome, making it difficult to identify the proper clone. Nonetheless, this procedure has been successfully used to clone most of the anthocyanin loci (Wienand and Saedler, 1987).

The C1 locus was the first regulatory locus in plants to be cloned, which was accomplished by transposon tagging (Paz-Ares et al., 1986; Cone et al., 1986). In both experiments, the transposable element En/Spm was used as a probe to clone En/Spm induced C1 mutants. These clones were confirmed to contain C1 specific segments by Southern analysis of various C1 mutants.

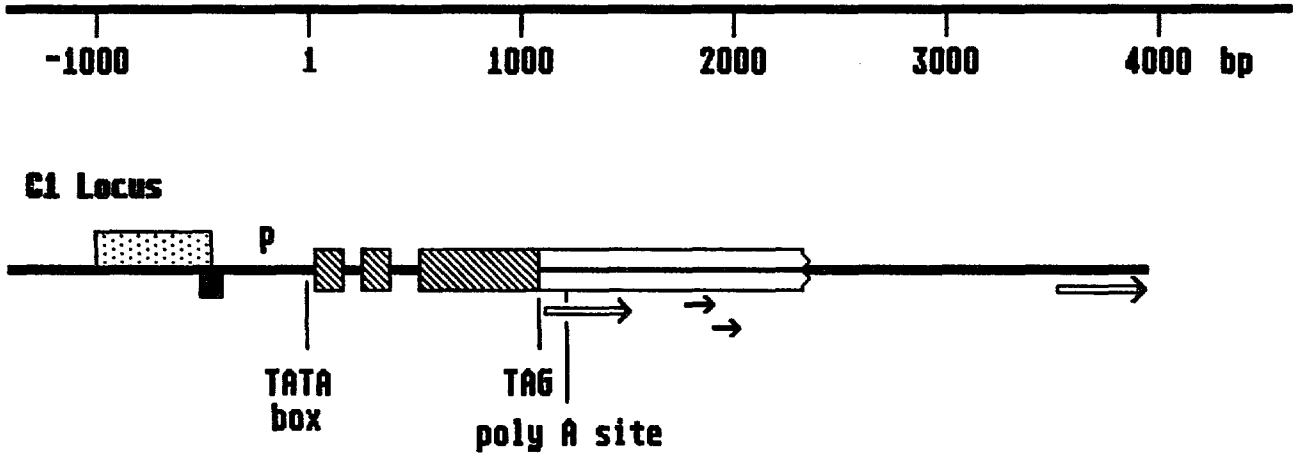
2.3.4. Molecular analysis of the wild type allele C1

The molecular analysis of C1 reveals that it has a rather complex nature (Paz-Ares et al. 1987). The gene codes for three transcripts (1.4, 1.6, and 2.5 kb) (Paz-Ares et al., 1986; Cone et al., 1986). Several cDNA libraries have been made to isolate the various C1 transcripts. For some unknown reason, these transcripts are very difficult, if not impossible, to clone (U. Wienand, Max-Planck-Institut für Züchtungsforschung, personal communication). To date, only one class of full length cDNAs have been isolated, 1.1 kb clones probably representing the 1.4 kb transcript (Paz-Ares et al., 1987). The difference in size of the cDNA clone and the transcript is probably because the measurement of the transcript was wrong. A partial clone of 2.1 kb with at least the poly(A) tail absent has also been isolated and probably corresponds to the 2.5 kb transcript.

Sequence analysis of the cDNA and genomic clones was conducted to elicit the structure of the C1 locus (Paz-Ares et al., 1987). C1 has 3 exons and two introns. Exons 1 and 2 are rather small, being 150 bp and 129 bp long. The two types of cDNA clones (1.1 and 2.1 kb) are identical for these two exons. Exon 3 has a variable length ranging from 720 bp for the 1.1 kb clones, and at least 1796 bp for the 2.1 kb clone. The two sets of clones share the same start site for transcription and the same reading frame. Therefore, they also have the same stop codon at position 1072. The C1 sequence is reported in Appendix. The CAAT box, TATA box, translation start site, stop codon and poly(A) addition sequence are highlighted in black. Introns 1 (88 bp) and 2 (144 bp) are given in small letters. A more diagrammatic depiction of the C1 locus is given in Figure 2.5.

The C1 locus has two sets of large direct duplications that are 90% homologous. The first set contains two copies that are approximately 400 bp. One copy is located from position 1088 to 1499 (Appendix) and the second extends from position 3535 to at least 3924 (Appendix), which is the end of the known sequence. The second set of duplications has two copies that are nearly 130 bp. Copy one starts at position 1754 and extends to 1882

Figure 2.5. The structure of the C1 gene. The box shaded with dots is the area of homology between C1 and C2. The boxes with lines represent the exons. The open box is the segment of exon 3 which is found in the 2.1 kb partial cDNA clone. "P" represents the promoter and "TAG" the translation stop site. The open arrows show the locations of the large direct duplications and the black arrows represent the smaller direct duplications

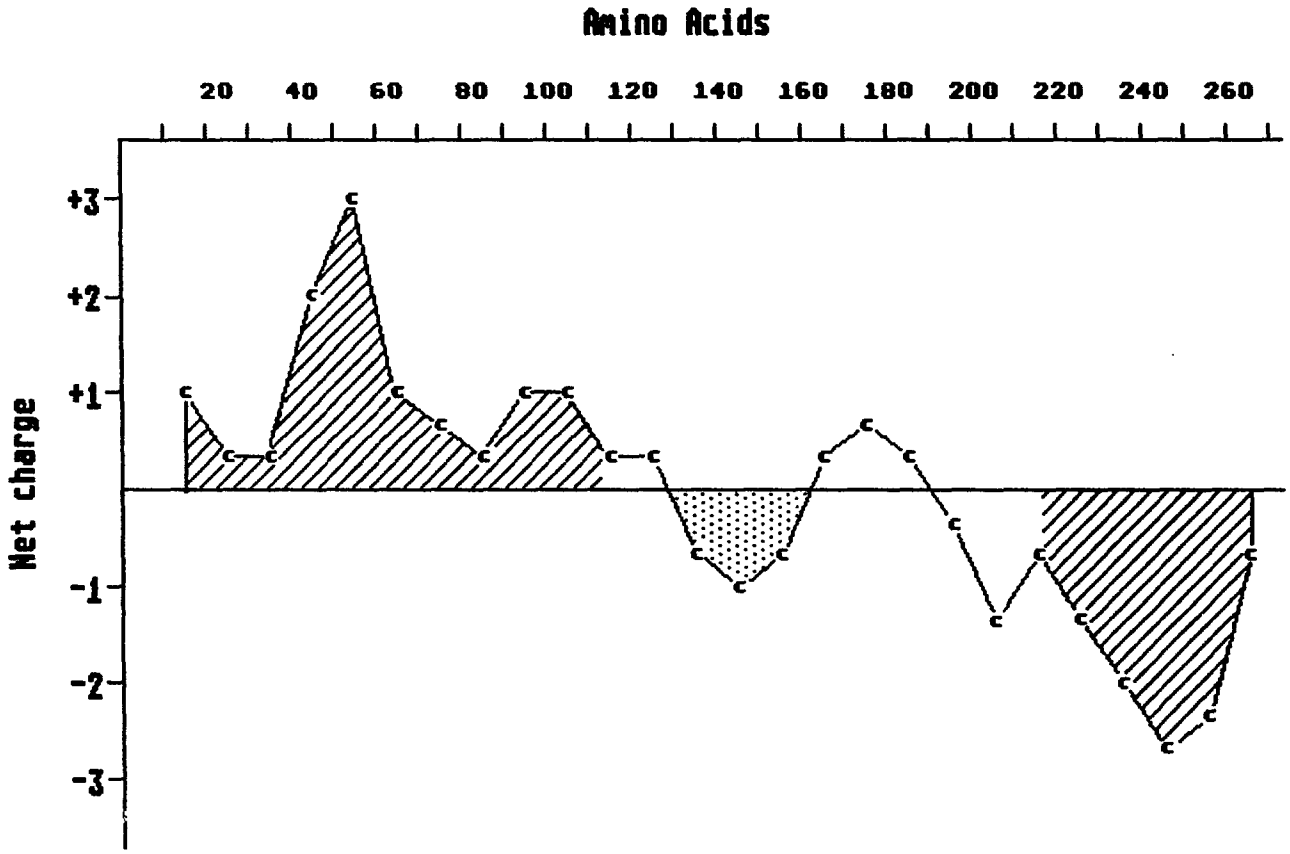

-1000
1
1000
2000
3000
4000 bp
C1 Locus
P
TATA
box
TAG
poly A site

(Appendix). The second copy overlaps with the first by 5 bp and starts at position 1877 and ends at 2005 (Appendix). The significance of these duplications has not been determined.

An additional sequence feature of C1 is that part of the sequence has homology to the C2 gene. A 550 bp region, from position -1013 to -465 (Appendix), has 80% homology to a region that is 1-1.5 kb upstream of the C2 gene (Figure 2.5). C2 codes for chalcone synthase (Figure 2.1) and is under control of the C1 locus (Dooner, 1983). Other structural genes tested in the anthocyanin pathway (A1 and Bz1) do not seem to have homology to this region. The function of this region is unknown, but it could be important in regulating gene expression.

The sequence of the putative C1 protein was derived from the theoretical translation of the DNA sequence of the cDNAs into the corresponding amino acids. The protein has 273 amino acids and a molecular weight of about 29 kd. There are two major domains in the protein. One is a basic domain that is located at the amino terminus between amino acids 1 and 114. The second is an acidic domain situated between amino acids 234 and 261 at the carboxy terminus (Figure 2.6). A minor acidic domain, that may be

Figure 2.6. A graph of the charge distribution of the putative C1 protein. This was done by averaging the net charge over 30 successive amino acids, measured at 10 amino acid intervals (lysine and arginine are basic, and glutamic acid and aspartic acid are considered acidic). The areas shaded with lines represent the major basic and acidic domains. The area shaded with dots indicates the minor acidic domain


Amino Acids
20
40
60
80
100
120
140
160
180
200
220
240
260
Net charge
+3
+2
+1
-1
-2
-3

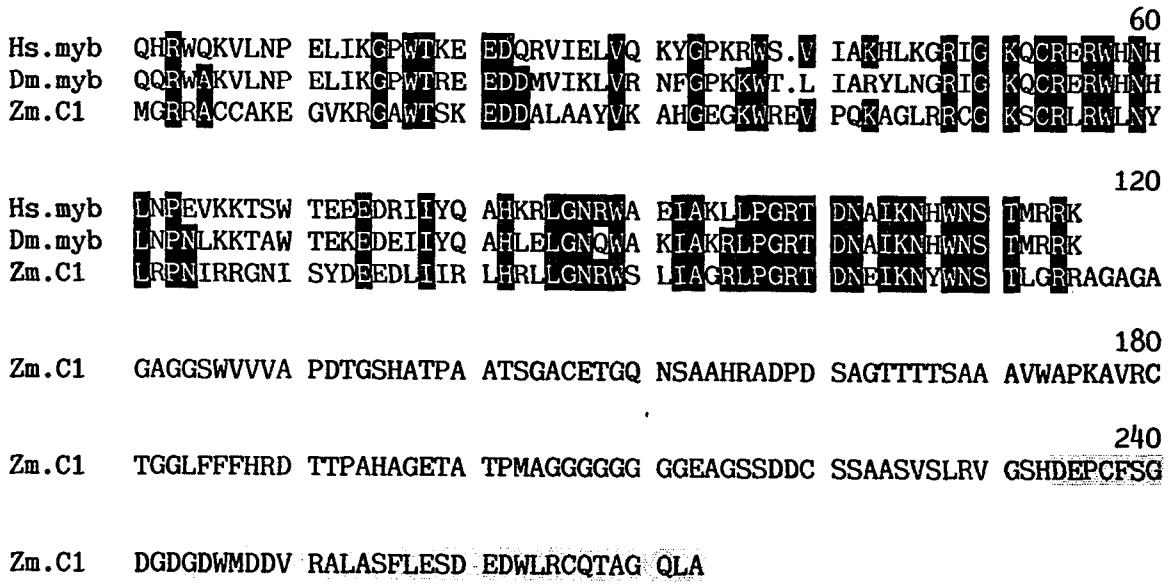
of importance, is located between amino acid positions 135 and 154 (Paz-Ares et al., 1987).

The sequence of the putative C1 protein was compared to other available sequences. Paz-Ares et al. (1987) discovered that the basic domain of C1 had 40% homology to the basic domains of myb proto-oncogene proteins (Figure 2.7). These myb proto-oncogene proteins are known to reside in the nucleus (Boyle et al., 1984) and that they are DNA binding proteins (Klempnauer and Sippel, 1987). Recently, it has been shown that the v-myb proto-oncogene protein binds to a DNA sequence that has AACG as the core sequence (Biedenkapp et al., 1988).

The region of homology between the C1 and myb proto-oncogene proteins (the basic domain) is also the DNA binding domain of the myb proto-oncogene proteins. This homology indicates that the C1 protein could also bind to DNA, and that it may govern anthocyanin biosynthesis by binding to cis-regulatory sequences of the structural anthocyanin loci, thereby regulating their expression (Paz-Ares et al., 1987).

Wienand et al. (1989) have shown that the C1 protein, expressed in E. coli, can bind to the A1 and C2 structural anthocyanin loci. The protein binds to the promoter of

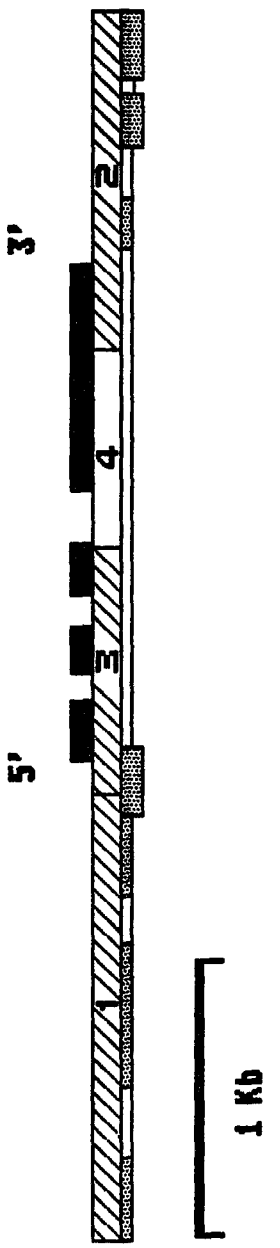
Figure 2.7. Sequence comparison of the putative C1 protein to the basic domains of myb proto-oncogenes from Homo sapiens and Drosophila melanogaster. Common amino acids are highlighted in black. The acidic domain of the C1 protein is shaded


60
Hs.myb QHRWQKVLNP ELIKGPWTKE EDQRVIELVQ KYGPKRWS.V IAKHLKGRIG KQCRERWHNH
Dm.myb QQRWAKVLNP ELIKGPWTRE EDDMVIKLVR NFGPKKWT.L IARYLNGRIG KQCRERWHNH
Zm.C1 MGRRACCAKE GVKRGAWTSK EDDALAAYVK AHGEGKWREV PQKAGLRRCG KSCRLRWLNY
120
Hs.myb LNPEVKKTSW TEEEDRIIYQ AHKRLGNRWA EIAKLLPGRT DNAIKNHWNS TMRRK
Dm.myb LNPNLKKTAW TEKEDEIIYQ AHLELGNQWA KIAKRLPGRT DNAIKNHWNS TMRRK
Zm.C1 LRPNIRRGNI SYDEEDLIIR LHRLLGNRWS LIAGRLPGRT DNEIKNYWNS TLGRRAGAGA
180
Zm.C1 GAGGSWVVVA PDTGSHATPA ATSGACETGQ NSAAHRADPD SAGTTTTSAA AVWAPKAVRC
240
Zm.C1 TGGLFFFHRD TTPAHAGETA TPMAGGGGGG GGEAGSSDDC SSAASVSLRV GSHDEPCFSG
Zm.C1 DGDGDWMDDV RALASFLESD EDWLRCQTAG QLA

the C2 gene and to several locations in the A1 gene. The major binding sites in the A1 locus are 5' and 3' to the coding region, and detailed analysis shows that the C1 protein probably interacts with the repetitive sequence AACGTTT (Figure 2.8). The binding site of the v-myb proto-oncogene protein has AACG as a core sequence (Biedenkapp et al., 1988) and this protein can also bind to the A1 gene, but the binding of the C1 protein is more sequence specific (U. Wienand, Max-Planck-Institut für Züchtungsforschung, personal communication).

There is an acidic domain with an amphipathic α-helical structure located at the carboxy terminus of the C1 protein (Figure 2.6). Similar short acidic domains are important components of the yeast transcriptional activators GCN4 and GAL4 (Struhl, 1987; Ptashne, 1988). These proteins bind to specific DNA sequences in the promoters of some important genes. GCN4 binds to some amino acid biosynthetic enzyme loci and GAL4 binds to galactose metabolizing (GAL) genes. The acidic domain of both proteins interacts directly or indirectly with the RNA polymerase II complex to activate transcription of the genes. The finding that the C1 protein has both acidic and DNA binding domains, strongly indicates that it acts as a transcriptional activator to regulate the expression

Figure 2.8. A diagram of the binding sites of the C1 protein expressed in E. coli to the A1 locus. The black boxes are the exons of the A1 locus. The numbered boxes represent DNA fragments that were used in filter binding assays. Those fragments that the C1 protein bound to are shaded with lines. The dotted boxes represent areas where the C1 protein binds as determined by DNA footprint analysis. Thick boxes indicate strong protection sites


5'
3'
1
2
3
4
1 Kb

of the structural anthocyanin genes (Paz-Ares et al., 1987).

2.3.5. Molecular analysis of the C1-I allele

The C1-I allele is dominant to all known C1 alleles. The production of anthocyanins in the aleurone layer is suppressed when C1-I is present. When kernels have two doses of the wild type allele and one of C1-I, the aleurone has light pigmentation. If a kernel has two or more copies of C1-I, there is no anthocyanin production. The molecular analysis of C1-I is important because it would help in understanding why some alleles are dominant to others, and how it is possible to have a dominant mutant with a negative phenotype.

Northern and Southern analyses demonstrated that C1-I had a major (1.3 kb) and a minor (1.45 kb) transcript, and that there is an insert in the gene (Cone et al., 1986). Paz-Ares et al. (1989) cloned the C1-I allele and its corresponding cDNAs. Heteroduplex and Southern analyses of the C1-I genomic clone compared with the wild type allele revealed that in the C1-I allele there was a 3.8 kb insertion and that at the same site or a site nearby the insertion, a 370 bp deletion occurred. Sequence and Northern analyses of the C1-I cDNAs showed that the minor transcript contained a 140 bp insert in contrast to the

major transcript, and this insertion corresponds to intron 2 of C1.

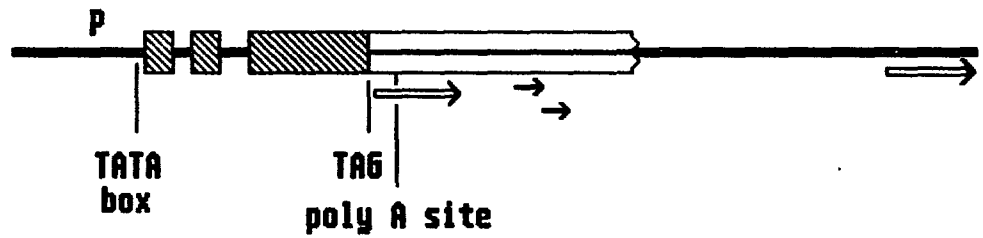
A sequence comparison of the genomic and cDNA clones of C1-I to C1 indicated there were several alterations in C1-I. One major alteration was that the 3.8 kb insert is located 8 bp 3' to the translation stop site (Figure 2.9), and because of this the two cDNAs extend about 200 bp into the insertion. Since intron 2 was not spliced out in the case of the 1.45 kb transcript this may indicate that the 3' region is important for normal RNA splicing and that the 3.8 kb insert prevents this. The most significant alteration in C1-I was that there is an 8 bp insertion at position 993 (Appendix). This small insertion is in the coding region and it causes a frameshift mutation (Figure 2.9).

The putative C1-I protein is 252 amino acids in length, which is much shorter than the wild type protein (273 amino acids), because the frameshift mutation leads to the premature termination of the open reading frame at position 1005 (Appendix). Due to this premature termination, the C1-I protein is missing the amino acids that constitute the major acidic domain (Figure 2.10). In addition, other base pair alterations make the minor acidic domain less acidic (Figure 2.11). These

Figure 2.9. A comparison of the structure of C1 and C1-I. The boxes with lines represent the exons. "P" represents the promoter. "TAG" and "TGA" are the translation stop sites. The open boxes represent segments of the transcripts which are not translated. The open arrows show the locations of the large direct duplications and the black arrows represent the smaller direct duplications


-1000
1
1000
2000
3000
4000 bp

C1 Locus


P
TATA
box
TAG
poly A site

C1-I

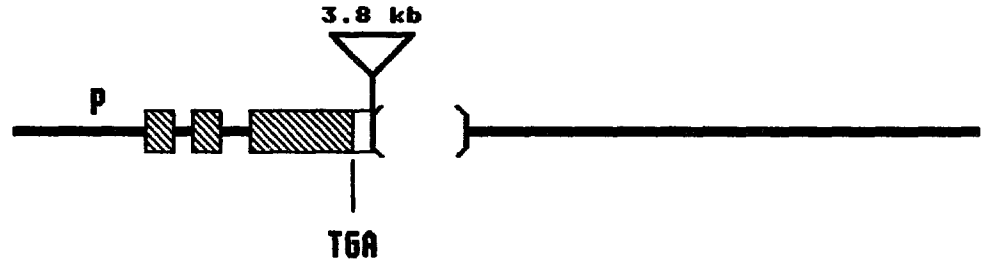

3.8 kb
P
TGA

Figure 2.10. Sequence comparison of the C1 and C1-I putative proteins. Differences in the amino acid sequence are highlighted in black. The acidic and basic domains are boxed in

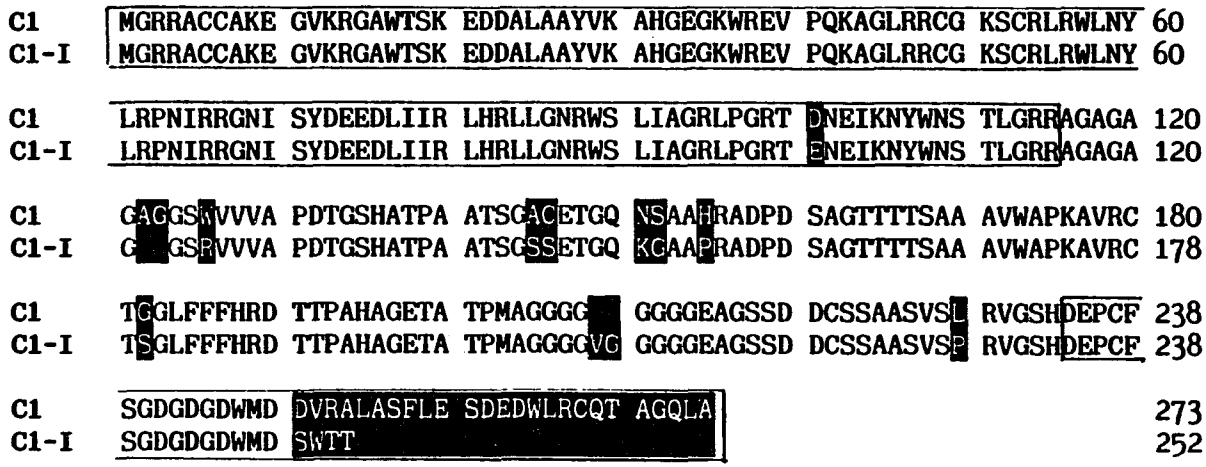
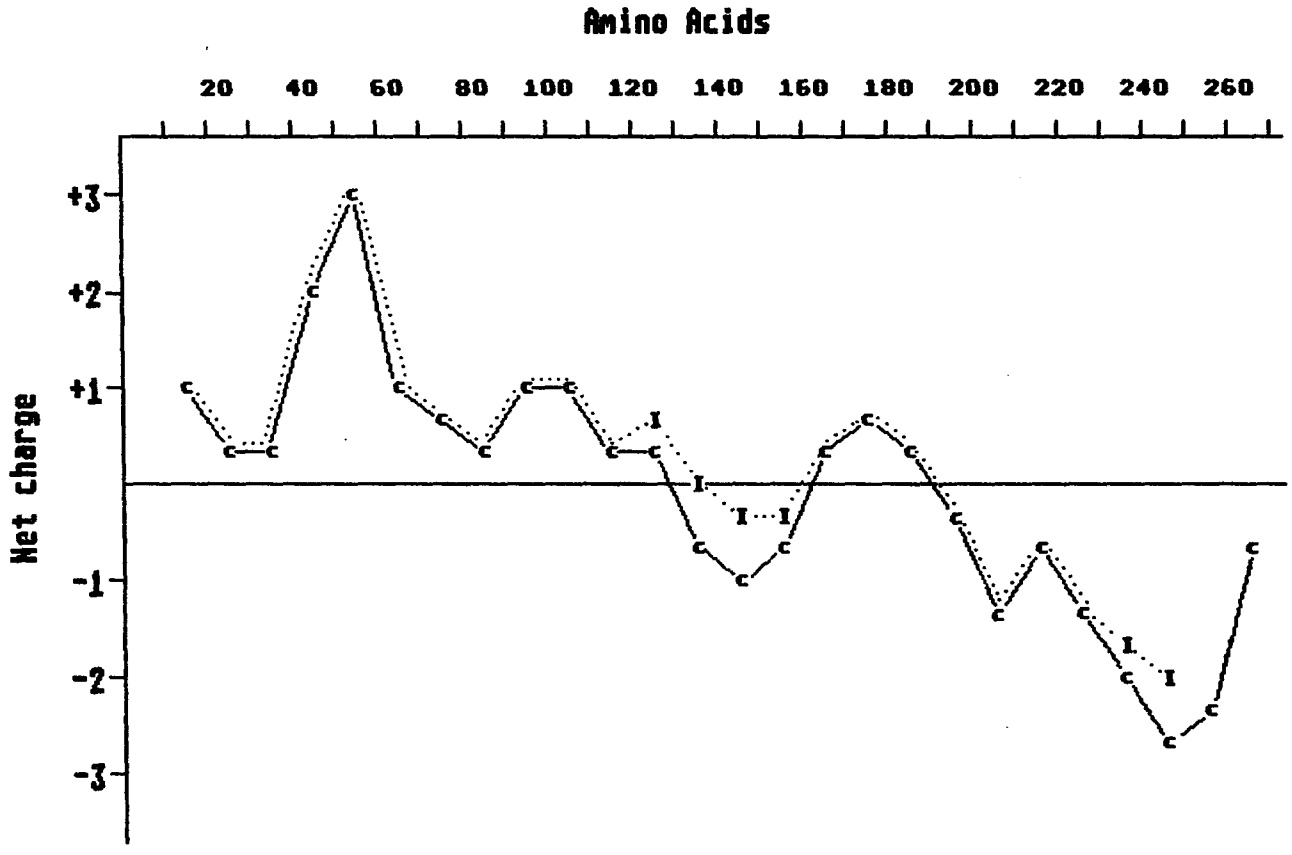

C1 MGRRACCAKE GVKRGAWTSK EDDALAAYVK AHGEGKWREV PQKAGLRRCG KSCRLRWLNY 60
C1-I MGRRACCAKE GVKRGAWTSK EDDALAAYVK AHGEGKWREV PQKAGLRRCG KSCRLRWLNY 60
C1 LRPNIRRGNI SYDEEDLIIR LHRLLGNRWS LIAGRLPGRT DNEIKNYWNS TLGRRAGAGA 120
C1-I LRPNIRRGNI SYDEEDLIIR LHRLLGNRWS LIAGRLPGRT ENEIKNYWNS TLGRRAGAGA 120
C1 GAGGSRVVVA PDTGSHATPA ATSGACETGQ NSAAHRADPD SAGTTTTSAA AVWAPKAVRC 180
C1-I G GSRVVVA PDTGSHATPA ATSGSSETGQ KCAAPRADPD SAGTTTTSAA AVWAPKAVRC 178
C1 TGGLFFFHRD TTPAHAGETA TPMAGGGG GGGGEAGSSD DCSSAASVSL RVGSHDEPCF 238
C1-I TSGLFFFHRD TTPAHAGETA TPMAGGGGVG GGGGEAGSSD DCSSAASVSP RVGSHDEPCF 238
C1 SGDGDGDWMD DVRALASFLE SDEDWLRCQT AGQLA 273
C1-I SGDGDGDWMD SWTT 252

Figure 2.11. A comparison of the charge distribution of the C1 and C1-I proteins from Figure 2.10. This was done by averaging the net charge over 30 successive amino acids, measured at 10 amino acid intervals (lysine and arginine are basic, and glutamic acid and aspartic acid are considered acidic). The C1 protein is plotted with a solid line and each data point is marked with a "c". The C1-I protein is plotted with a dotted line and the data points are marked with an "I" unless they share the same location with the C1 protein and then only "c" is shown


Amino Acids
20
40
60
80
100
120
140
160
180
200
220
240
260
Net charge
+3
+2
+1
-1
-2
-3

alterations in the acidic domains indicate that the C1-I protein would not have a transcriptional activator function. In fact, it has been shown that when the acidic domain of the GAL4 transcriptional activator is eliminated, the protein is converted into a transcriptional repressor (Hope et al., 1988). Because of these similarities, and the fact that the DNA binding domain of the C1-I protein is intact, Paz-Ares et al. (1989) concluded that the C1-I protein would probably also act as a transcriptional repressor.

The dominant nature of C1-I can probably be attributed to the fact that it acts as a transcriptional suppressor. The C1-I protein is capable of binding to DNA (U. Wienand, Max-Planck-Institut für Züchtungsforschung, personal communication), but since it has no activator domain it would block transcription. The A1 gene is known to have several locations where the C1 protein can bind (Figure 2.8). If the C1-I protein is residing at any of these sites it could be sufficient to block transcription, even though C1 could be bound to an adjacent site. Thus, C1-I is dominant to C1 in the heterozygous condition. However, if large amounts of C1 protein are present, such as in C1 C1/C1-I kernels, the inhibiting effect of C1-I can be overcome (Paz-Ares et al., 1989).

2.4. Light Induction of Anthocyanin Loci

Phenotypically, several anthocyanin loci are known to be inducible by light. Often this phenomenon is limited to a specific allele for a locus. c1-p and pl are examples of light inducible alleles. The Pl locus is a duplicate gene of C1, except that it is expressed in plant tissue. Plants homozygous for pl and containing all the other loci necessary for color formation, produce anthocyanins when the tissue is exposed to light.

Little is known about how light induces the production of anthocyanins. The best documented example is the chalcone synthase gene in parsley (Petroselinum crispum) (Dangl et al., 1989). The major light stimulus for the CHS gene in parsley is ultraviolet light (UV) (Wellmann, 1975), although blue and red light (phytochrome) can have an influence. Short term (2-4 hours) or continuous UV irradiation of dark-grown parsley suspension cells results in a rapid increase of CHS enzyme activity, mRNA level, and transcriptional activity, as measured by run-off transcription experiments on isolated nuclei (Schröder et al., 1979; Kreuzaler et al., 1983; Chappell and Hahlbrock, 1984).

In order to understand the molecular basis of how UV irradiation stimulates CHS expression, Schulze-Lefert et

al. (1989) analyzed the promoter for cis-regulatory sequences via *in vivo* DNA footprinting. Dark grown cultured parsley cells were treated with UV irradiation, followed by chemical degradation sequencing for the G nucleotide. Footprint analysis was done by the procedure of Church and Gilbert (1984). Three light dependent footprints were found. Deletion analysis of the promoter and mutagenesis of the DNA sequence within the footprint boxes demonstrated that box I (TGAGCTCATG) and II (AGGTACCTAG) were necessary for light response.

Sequence comparisons of boxes I and II to known genes revealed that there were no significant homologies to box I, but box II had a high degree of homology to several genes. The composition and position of box II is conserved in CHS loci of other plants. Although the promoters of *C2* and *Whp* have little homology to each other, they both contain box II. There is no box II in the *C1* locus. Box II is also found in the ribulose bisphosphate carboxylase small subunit genes of numerous plant species (Schulze-Lefert et al., 1989). Because box II is also located in genes under pathogenic (*rolbc*) and anaerobic (*adh*) control, it indicates that UV induction is not the only stimulus that can induce the production of proteins that bind to box II. Schulze-Lefert et al.

(1989) proposed that box II may be employed by other stress-induced responses, in addition to light regulation. Further analysis is needed to determine the interaction between boxes I, II, and III, in addition to defining the proteins binding to them.

2.5. Eukaryotic Transcriptional Activators

Transcriptional activators are specific proteins that bind to DNA and can activate the transcription of a wide variety of genes. These proteins have two important components. One is the DNA binding domain and the other is the activating domain. The binding domain interacts only with a specific DNA sequence and therefore, protein sequence alterations in this area are tolerated to a limited extent (Ptashne, 1988).

The only constructional restraints in the active domain are that it must be acidic and have an amphipathic α-helical structure. The α-helices bear hydrophobic residues upon one surface and negatively charged particles on the other. Besides these structural features, there is no similarity in the protein sequence of different activator domains. Giniger and Ptashne (1987) demonstrated this by replacing the active domain of GAL4 with two oligonucleotides coding for the same 15 amino-acid peptide. One peptide was constructed in a manner

that it could form an amphipathic α-helical structure. In the other peptide, the amino acids were identical to the first peptide, but they were in a different order, so the resulting configuration was not an amphipathic α-helical structure. When tested, only the peptide with the amphipathic α-helical structure behaved as an activator.

The active and DNA binding domains are independent of each other and interchangeable with a corresponding domain from another transcriptional activator (Brent and Ptashne, 1985). A hybrid protein with the GAL4 active domain fused to the DNA binding domain of the LexA repressor will not initiate transcription of the GAL genes, because there is no binding site for the LexA repressor protein. When the GAL4 DNA binding site in the promoter of a GAL gene is replaced with the binding sites for the LexA repressor, the fusion protein can bind, and will activate the gene (Brent and Ptashne, 1985).

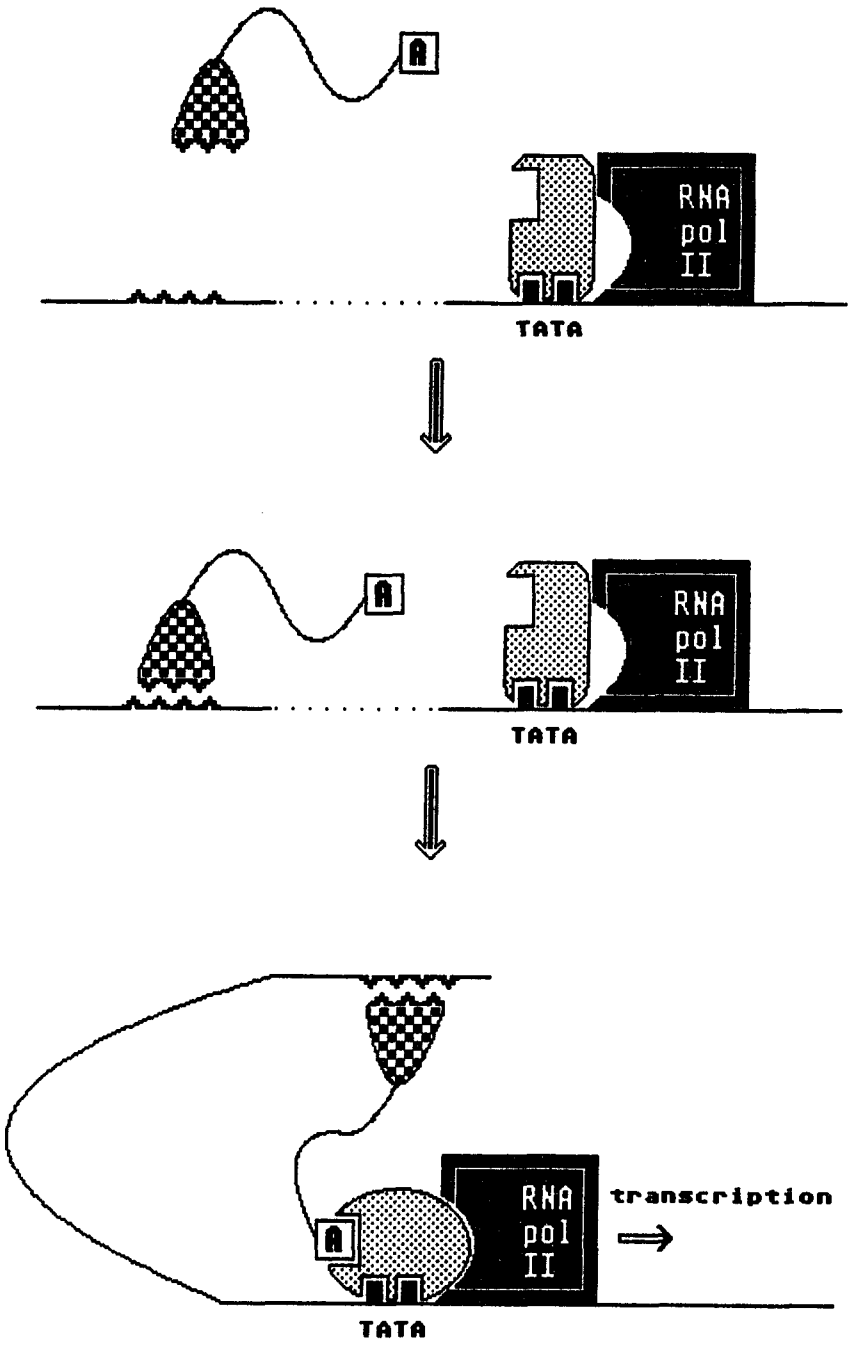
Numerous studies have been conducted by exchanging or altering the acidic and basic domains, and a brief summary of the findings is given below. For more extensive reviews see Struhl (1987) and Ptashne (1988). An increase in acidity in the activator domain, without destroying the amphipathic α-helical structure, augments the activation function (Gill and Ptashne, 1987, 1988). If an activator

domain is very potent, a high amount of acidity, as in the case of the herpes simplex virus protein Vp16, the DNA site where the protein binds can be located at a distance further away from the gene than a weak activator domain (Sadowski et al., 1988). Therefore, transcriptional activators could have binding sites in enhancer elements, which are usually located several hundreds of base pairs 3' or 5' to of the gene. In this case, the gene would be activated from a distant location (Ptashne, 1988).

It is not known how the active domain functions. The active domain probably does not associate directly with RNA polymerase II, but could interact with another protein that has this ability, such as the TATA binding protein (Ptashne, 1988). In this case, the transcriptional activator would bind to its predetermined site and then interact with the TATA binding protein. This protein is bound to the TATA site, and can interact with RNA polymerase II due to its association with the transcriptional activator (Figure 2.12).

This example is a simplified model. Usually several proteins are involved in regulating the transcription of a gene. Transcriptional activators can act in a synergistic fashion. GAL4 and the mammalian transcriptional activators ATF (activating transcription factor) and USF

Figure 2.12. A model showing how a transcriptional activator may work. The transcriptional activator (the checkered segment is the DNA binding domain and the acidic domain is labelled with an "A") binds to the DNA. The acidic domain interacts with the TATA binding protein, and when this happens the TATA binding protein can stimulate RNA polymerase II to transcribe the gene. If the binding site of the transcriptional activator is far away from the TATA binding protein, the acidic domain is brought in proximity by looping out the intervening DNA


A
RNA
pol
II
TATA
A
RNA
pol
II
TATA
A
RNA
pol
II
transcription
TATA

(upstream stimulating factor) individually enhance transcription by a factor of two or four-fold. In contrast, when these proteins complement each other, they stimulate transcription by more than fifty-fold (Lin et al., 1988).

Transcriptional activators themselves must be regulated. This can be done by limiting their expression or a second protein can bind to them preventing them from carrying out their designated function. The GAL4 protein, which activates genes involved in galactose catabolism in yeast, is inhibited by the protein GAL80. GAL80 binds to the activator domain, but not the binding domain of GAL4. In the presence of galactose the GAL80 protein is released from GAL4 and galactose metabolism is turned on (Johnston et al., 1987; Lue et al., 1987; Ma and Ptashne, 1987). If a transcriptional activator was not regulated, it could continuously stimulate the transcription of a particular gene. Accumulation of the gene's product could be lethal to the organism.

2.6. The Formation of mRNA and its Role in Gene Expression

Important regions regulating transcription are relatively easy to define. This can be done either by DNA footprint analysis or expression studies with altered

promoters. In contrast, analysis of the processing of RNA is rather difficult. The complication arises because several processing steps may occur while the primary RNA is being transcribed and therefore, intermediate products are difficult, if not impossible, to isolate (Levitt et al., 1989; Proudfoot, 1989). To achieve a mature mRNA, a cap structure (7MeGppp) is added to the 5' end of the transcript, the introns are spliced out and, in most instances, the 3' end is polyadenylated [poly(A+)]. In addition to these processing steps, transcription via RNA polymerase II must be terminated. Termination normally occurs near the poly(A) addition signal, but it can also happen several kilobases down from this site (Hagenbüchle et al., 1984). Unlike prokaryotes, termination of transcription in eukaryotes seems to be a somewhat random process because the mature mRNA is generated from a rather heterogenous pool of transcripts, differing in the length of their 3' ends. However, it does seem that pause sites, poly(A) signal (AATAAA) and a GT-rich region downstream of the poly(A) signal, which is also important for polyadenylation, play important roles in terminating transcription and the amount of mature mRNA produced (Connelly and Manley, 1988; Proudfoot, 1989).

Evidence that the poly(A) signal is important in termination comes from the study of the human α2 globin gene and a mutant allele that has an altered poly(A) signal (AATAAG) (Higgs et al., 1983; Whitelaw and Proudfoot, 1986). Transcription of the mutant allele reads past the normal termination site and a poly(A) signal that is 1.kb 3' of the normal signal is used instead of the mutant signal. In addition, the mutant allele produces only 20% of the normal level of mRNA. The only major structural difference between the wild type α2 allele and the mutant is the base pair change of A to G at the end of the signal sequence. This study showed a functional poly(A) signal is needed for 3' processing and transcription termination.

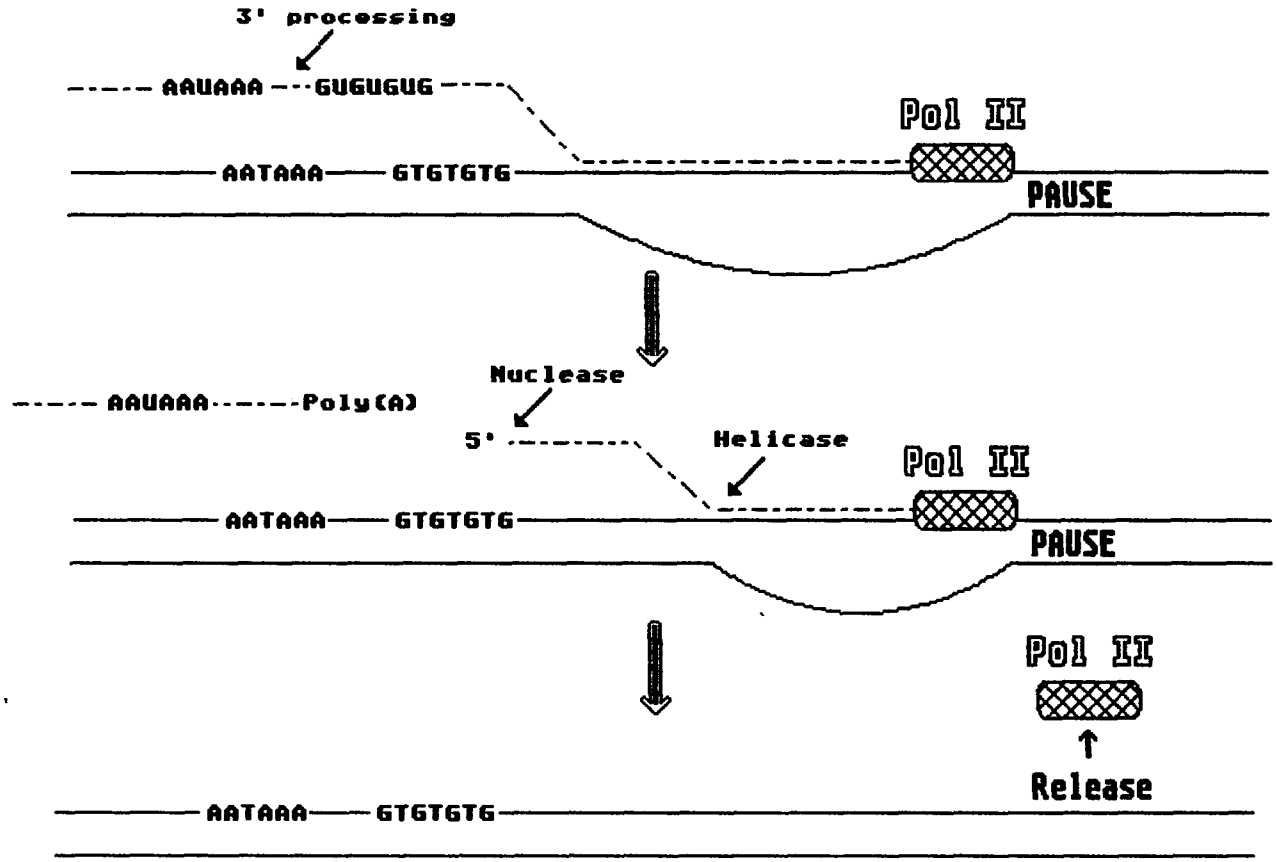
A GT-rich region 3' of the poly(A) signal is also important for transcriptional termination, at least in mammalian systems (Proudfoot and Whitelaw, 1988). The G-T rich region and the poly(A) signal constitute a poly(A) complex. Deletion of the GT-rich region or the poly(A) signal of the SV40 early gene abolishes cytoplasmic mRNA production, but the level of nuclear RNA is substantially higher than when the normal polyaddition complex is present. Analysis of the nuclear RNA shows that the transcripts are very large and are produced by RNA

polymerase II transcribing several times around the plasmid construct. These large transcripts are produced because transcription is not terminated (Connelly and Manley, 1988).

Large RNA transcripts are also observed in a mutant form of the DNA tumor virus polyoma. In this case the polyoma gene has a weak poly(A) complex (Acheson, 1984). Replacement of the weak poly(A) complex with a strong one (Gil and Proudfoot, 1987) prevents the formation of these long concatamers and terminates transcription slightly beyond the new poly(A) complex (Lanoix and Acheson, 1988). Both examples (SV40 and polyoma) show that RNA polymerase II continues to transcribe unless an efficient poly(A) complex is available for 3' processing and transcription termination (Proudfoot, 1989).

It is believed that active 3' processing is an important factor in terminating transcription. Although several models have been proposed for the termination mechanism, one is considered the most likely scenario (Connelly and Manley, 1988; Proudfoot and Whitelaw, 1988; Proudfoot, 1989). In this model (Figure 2.13), the incipient mRNA is released from the remaining transcript and a poly(A) tail is added as soon as RNA polymerase II passes a functional poly(A) complex. Meanwhile, RNA

Figure 2.13. A model of the termination of RNA polymerase II. RNA polymerase II (Pol II) may be slowed by a pause site. The 3' processing mechanism cleaves the RNA and adds the poly(A) tail. The nascent RNA transcript is not protected by a cap structure at the 5' end and a 5'->3' exonuclease degrades the RNA. Meanwhile, a helicase unravels the DNA:RNA hybrid. When the transcript is degraded, RNA polymerase II is released from the DNA template


3' processing
AAUAAA
GUGUGUG
Pol II
AATAAA
GTGTGTG
PAUSE
Nuclease
AAUAAA
Poly(A)
5'
Helicase
Pol II
AATAAA
GTGTGTG
PAUSE
Pol II
Release
AATAAA
GTGTGTG

polymerase II continues to transcribe downstream. The transcript, unlike the nascent mRNA, does not have a protective 5' cap structure and thus is susceptible exonuclease activity. A 5' -> 3' exonuclease degrades the transcript while a helicase unwinds the RNA transcript from the DNA template. When exonuclease has fully degraded the transcript, RNA polymerase II is released from the DNA template. Therefore, termination is dependent on several factors and for this reason there is usually not an exact site where termination occurs.

If a gene were to have an inefficient poly(A) complex, RNA polymerase II may travel a considerable distance downstream before the nascent mRNA is cut out. The exonuclease and helicase would have more transcript to work on before catching up with the polymerase. This also would lead to heterogeneity of termination.

Pause sites (Figure 2.13) are also believed to be present in the 3' flanking regions and would affect the location of termination. Pause sites would slow RNA polymerase II down and give the helicase and exonuclease time to catch up (Connelly and Manley, 1988; Proudfoot, 1989). Several possible mechanisms could stimulate pausing. Inverted repeats can promote the formation of hairpin loops that cause the RNA polymerase to pause

(Pribyl and Martinson, 1988). RNA polymerase also can be stalled by a protein binding to the DNA template. The removal of the CAAT box in the major late protein-promoter of SV40 allows read-through transcription. In contrast, when the CAAT box and the protein that binds to it are present, transcription stops near the CAAT box (Proudfoot, 1989).

It is possible to imagine that a 3' flanking region may contain several pause sites. A fraction of the transcripts would be terminated at each site, which would create a heterogenous termination. As proposed by this model (Figure 2.13) termination requires a stop site and a poly(A) complex.

Almost all studies on 3' processing and transcription termination have been conducted in mammalian systems. Little is known about these functions in plants. Plants have a poly(A) signal, but it is not as highly conserved as in mammals (Heidecker and Messing, 1986; Walbot and Messing, 1988). In several instances, the terminal nucleotide in plants is G instead of A (AATAAG). No study has been conducted to determine if a G-T rich region 3' to the poly(A) signal is present and necessary in plant genes. Plant genes seem to have more than one poly(A) site, but the reason for this is not known (Heidecker and

Messing, 1986; Walbot and Messing, 1988). Perhaps several plant genes have weak poly(A) complexes and therefore, polyadenylation can occur at more than one site.

3. MATERIAL AND METHODS

3.1. Abbreviations

Abbreviation	Term
ATP	Adenosine triphosphate
bp	base pair(s)
BSA	Bovine serum albumin
β-ME	β-mercaptoethanol
cpm	counts per minute
DEAE	Diethyl pyrocarbonate
Dieca	Na-diethyldithiocarbaminate
DMS	Dimethylsulfate
dNTP	Deoxynucleoside triphosphate
EDTA	Ethylenediaminetetraacetate
IPTG	Isopropyl-β-D-thiogalactoside
kb	kilobase(s)
OD	Optical density
PNK	Polynucleotide kinase
RNAse	Ribonuclease
SDS	Sodium dodecyl sulphate
Tris	Tris (hydroxymethyl) aminomethane
TE	Tris-EDTA buffer (10 mM Tris pH 7.5, 1 mM EDTA)
UV	Ultraviolet
X-gal	5-bromo-4-chloro-3-indolyl-B-Dgalactoside

One letter code for amino acids

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic acid
C	Cysteine
Q	Glutamine
E	Glutamic acid
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine

3.2. Isolation of Maize Nucleic Acids

3.2.1. mRNA extraction

The following protocol was provided by Dr. Andrew Hudson (Cambridge University, Cambridge, England).

1. Freeze 15 g of tissue in liquid nitrogen and grind to a fine powder. Put tissue in a disposable 50 ml centrifuge tube.
2. Thaw tissue in 20 ml of extraction buffer and 5 ml of 10% SDS.
3. Shake mixture.
4. Top the tube off with phenol and shake briefly.
5. Centrifuge for 5 min. at 6000 r.p.m.
6. Transfer upper phase to clean centrifuge tube.
7. Repeat steps 4-6 two times using a phenol:chloroform (1:1 mixture) extraction and one time with a chloroform extraction.
8. Add 16 ml of 5 M LiCl.
9. Mix and precipitate RNA overnight at 4°C. Precipitated RNA will appear white and fluffy.
10. Centrifuge at 6000 r.p.m for 10 min. at 4°C.
11. Resuspend pellet in oligo (dT) binding buffer for Poly (A)+ mRNA extraction, without the SDS. For total RNA, resuspend in 0.2 M LiCl and precipitate in two and half volumes of EtOH.

12. Treat 0.5 grams of oligo (dT)-cellulose with a solution of 0.1 M NaOH and 5 mM EDTA.
13. Equilibrate the oligo (dT)-cellulose with the binding buffer.
14. Place the oligo (dT)-cellulose in a sterile 50 ml Falcon tube. After the removal of the binding buffer, add the RNA sample. Place the tube on a shaker to gently mix the sample for 15 min.
15. Centrifuge the sample for 2 min. at 2000 r.p.m. and pour off the supernatant.
16. Wash the sample with 25 ml of the binding buffer.
17. Repeat steps 15 and 16 two more times.
18. Resuspend the cellulose in 25 ml of binding buffer and pour the cellulose into a column.
19. Wash the column with 50 ml of the binding buffer.
20. Release the mRNA from the cellulose by adding the elution buffer. Regulate the flow of the buffer so that one drop falls every 10 seconds. Collect 0.5 ml samples in a quartz cuvette and take an OD reading. The cuvette is made RNase free by treating it with a solution of iodoacetic acid (313 mg in 100 ml H2O) for one hour. Collect all the samples and add 1/10 volume of sodium acetate (3 M, pH 5.2).

21. Precipitate the mRNA by adding 2.2 volumes of absolute EtOH and placing the sample in a -20°C freezer overnight.
22. Centrifuge the sample at 10,000 r.p.m. for 30 min. at 4°C.
23. Wash the mRNA pellet with 70% EtOH.
24. Store mRNA in 70% EtOH at -70°C until needed.

Extraction Buffer
50 mM Tris-HCl (pH 9.0)
150 mM NaCl or LiCl
5 mM EDTA

mRNA Binding Buffer
0.5 M NaCL
10 mM Tris-HCl pH 7.5
1.0 mM EDTA
0.5% SDS

mRNA Elution Buffer
10 mM Tris-HCl pH 7.5
1.0 mM EDTA
0.2% SDS

3.2.2. Plant genomic DNA extraction

The procedure followed was essentially as described by Schwarz-Sommer et al. (1984).

1. Chill DNA extraction buffer on ice.
2. Freeze desired amount of tissue in liquid nitrogen.
3. Grind tissue in a coffee grinder. Then place powder in extraction buffer, and shake gently. See buffer chart to determine the appropriate buffer:tissue ratio.
4. Place solution on ice for 60 to 90 min. and shake gently every 15 min.
5. Centrifuge in a swing-out rotor at 5000 r.p.m. for 20 min. at 5°C.
6. Pour supernatant through sterile cheese cloth into a sterile bottle.
7. Add an equal volume of phenol:chloroform. The phases are mixed by continuous gentle shaking for 5 min. followed by a 5 min. stationary period.
8. Separate the two phases by centrifugation in a swing-out rotor at 6000 r.p.m. for 10 min. at 10°C.
9. Transfer the upper phase to a sterile bottle using the wide end of a sterile pipette.

10. Two times the volume of EtOH is added slowly, and then mixed gently so that the DNA forms a fibrous mass.
11. The clump of DNA is transferred to a 50 ml Falcon tube using a sterile pipette. The DNA is washed with 70% EtOH and gently centrifuged. The 70% EtOH is poured off and the DNA is resuspended in 20 ml TE buffer overnight in the refrigerator.
12. Add 200 µl of 2 mg/ml RNase which has been pretreated for 2-3 min. at 80°C. The combined solution is heated at 37°C for 30 min.
13. Add 400 µl 10% SDS and 5 mg Proteinase K (5 mg suspended in 100 µl TE buffer) and shake lightly. Heat for 30 min. at 37°C, followed by 10 min. at 50°C.
14. Add an equal volume of phenol:chloroform and shake gently for 5 min. Separate phases by centrifugation at 10°C for 10 min. at 6000 r.p.m. Transfer upper phase to a new 50 ml Falcon tube using the wide end of a sterile pipette.
15. Repeat step 14 two times, except replace the phenol:chloroform with chloroform.

16. Add two times the volume of EtOH, mix, and place in -70°C for 30 min. Centrifuge at 6000 r.p.m. for 20 min. at 4°C. Wash pellet with 70% EtOH.

17. Dissolve DNA in 2 ml TE buffer. Add 500 µl EtBr and enough CsCl to reach a refractive index of 3.86.

18. Centrifuge at 40,000 r.p.m. at 15°C for 17 hours.

19. Remove DNA band and place into a sterile tube. Add one volume of TE buffer and mix.

20. Extract EtBr by adding one volume of 1-N-butanol saturated with CsCl and shaking gently. Remove the upper phase and add one volume of CsCl saturated 1-N-butanol. Repeat this procedure until the lower phase is colorless.

21. Precipitate DNA by adding three volumes of H2O, 8 volumes absolute EtOH and mix gently. Place the solution in -70°C for 30 min.

22. Centrifuge DNA at 8000 r.p.m. at 4°C for 30 min.

23. Dissolve pellet in appropriate amount of TE buffer. Fifteen grams of tissue will yield about 500 µg of DNA.

Genomic DNA Extraction Buffer
For 10 g of plant tissue:
30 ml of Dieca buffer
855 µl of Sarkosyl (37%)
16 µl of β-Mercaptoethanol
0.68 g of Dieca

Dieca buffer
0.1 M NaCl
50 mM EDTA
50 mM Tris-HCl (pH 7.6)

3.3. Handling DNA Fragments

3.3.1. Isolation of DNA fragments

The following protocol was derived from Maniatis et al. (1984).

1. Digest the desired DNA with the appropriate restriction enzyme according to the manufacturers' instructions.
2. Add loading buffer to digested DNA and load sample onto an agarose gel. The percentage of agarose and the length of run can be determined

from the information provided by Maniatis et al. (1984).

3. Stain the gel in a EtBr bath.
4. Under a low power UV lamp, cutout the desired fragment with a scalpel.
5. Place the fragment in a dialysis tube with agarose buffer.
6. Elude the fragment from the gel in the dark.
7. Remove buffer and extract DNA according to Section 3.3.3.

10X Agar Buffer (5 liters)
242 g Tris
20.5 g sodium acetate
18.5 g EDTA
pH 7.8 (use acetic acid)

Loading buffer (20 ml)
1 ml 1 M Tris
1 ml 0.4 m EDTA
10 ml glycerin
0.1% bromophenol blue

3.3.2. Preparation of dialysis tubes

1. Cut dialysis tubes to the desired length. Place tubing in a beaker and cover with a 0.5% solution of SDS.
2. Boil the tubes for 30 min.
3. Wash with H2O.
4. Boil tubing for 30 min. in 20 mM EDTA.
5. Wash two times with H2O.
6. Store tubing at 4°C in 1 mM EDTA.

3.3.3. Purification of DNA fragments

1. A DEAE or RPC5 column is washed with the elution buffer (1 M NaCl in TE) and then the binding buffer (0.25 M NaCl in TE).
2. DNA solution is added to a DEAE cellulose or RPC5 column.
3. Solutions in the RPC5 column can be forced through, while only gravity can be used with a DEAE cellulose column.
4. The column is washed with 500 µl binding buffer.
5. The DNA is released from the column using 100-500 µl of the elution buffer.
6. The DNA is precipitated with 2 volumes EtOH for 30 min. at room temperature. Centrifuge at 13,000 r.p.m. for 30 min. at room temperature.

7. The DNA is washed with 70% EtOH and then resuspended in the desired amount of TE buffer.

3.4. Working with Plasmids

3.4.1. Plasmid miniprep method

The plasmid miniprep method is a modification of the alkaline method reported in Maniatis et al. (1984).

1. Inoculate 5 ml of L-medium containing the proper antibiotic with a single bacterial colony. The culture is aerated by shaking and incubated at 37°C overnight.
2. A 1.5 ml aliquot is centrifuged in an Eppendorf tube at 13,000 r.p.m. for 2 min.
3. Remove the medium by aspiration and resuspend the pellet in 80 µl of ice cold 50 mM glucose, 10 mM EDTA, and 25 mM Tris-HCl pH 8.0. by vortexing.
4. Add 20 µl of 20 mg/ml lysozyme that has been dissolved in the buffer from step 3 and vortex briefly. Let tube stand for 5 min. at room temperature.
5. Add 200 µl of freshly prepared 0.2 N NaOH and 1% SDS. Invert the tube several times (do not vortex). Place the tube in ice for 5 min.

6. Add 150 µl of 3 M KAc pH 4.8 and vortex the tube in an inverted position for 10 sec. Place the tube in ice for 5 min.
7. Centrifuge tube at 13,000 r.p.m. at 4°C for 10 min.
8. Place supernatant into a fresh Eppendorf tube. Add an equal volume of phenol:chloroform and vortex for 2 min.
9. Centrifuge the tube for 2 min. at 13,000 r.p.m. and place the aqueous phase into a new tube.
10. Repeat steps 8 and 9 using a chloroform extraction instead of phenol:chloroform.
11. Add an equal volume of isopropanol, vortex and let the tube stand for 2 min. at room temperature to precipitate DNA. Centrifuge for 5 min. at 13,000 r.p.m. and remove supernatant with a drawn out Pasteur pipette.
12. Wash DNA with 70% EtOH. Remove supernatant with a drawn out Pasteur pipette and air dry for 2 min.
13. Resuspend DNA in 50 µl of TE with 50 µg/ml RNase.
14. Digest a 5-10 µl aliquot (5 µl for puc19 and 10 µl for PBR322) with the appropriate restriction enzyme.

3.4.2. Large scale LiCl plasmid preparation

1. Inoculate 200 ml of L-medium (containing the proper antibiotic) using a single bacterial colony. The culture is aerated by shaking and incubated at 37°C overnight.
2. Centrifuge the culture for 10 min. at 5000 r.p.m.
3. Resuspend the pellet in 20 ml of 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0, and 100 mg lysozyme.
4. Place on ice for 30 min.
5. Add 50 ml of freshly prepared 0.2 N NaOH and 1% SDS. Place on ice for 10 min.
6. Add 37.5 ml of 3 M KAc pH 4.8. Place on ice for 15 min.
7. Centrifuge at 0°C for 30 min. at 8000 r.p.m.
8. Filter supernatant.
9. Add 1 volume of isopropanol. Place on ice for 15 min.
10. Centrifuge at 20 °C for 30 min. at 4500 r.p.m.
11. Resuspend pellet in 6 ml of TE buffer. Add 6 ml LiCl and place on ice for 15 min.
12. Centrifuge at 20°C for 20 min. at 10,000 r.p.m.
13. Transfer the solution to a new tube and add one volume of isopropanol.

14. Place on ice for 15 min.
15. Centrifuge at 20°C for 30 min. at 13,000 r.p.m.
16. Resuspend in 3 ml of TE buffer with 60 µg RNase.
17. Incubate for 30 min. at 37°C.
18. Two times phenol:chloroform extraction.
19. One chloroform extraction.
20. Add 1/10 volume 3 M NaAcetate.
21. Precipitate DNA in two volumes EtOH at room temperature for 10 min.
22. Centrifuge at room temperature for 20 min. at 8,000 r.p.m.
23. Resuspend pellet in 500 µl TE buffer.

3.4.3. Small scale LiCl plasmid preparation

1. Inoculate 10 ml of L-medium (containing the proper antibiotic) using a single bacterial colony. The culture is aerated by shaking and incubated at 37°C overnight.
2. Centrifuge the culture for 10 min. at 6000 r.p.m.
3. Resuspend the pellet in 800 µl of 50 mM glucose, 10 mM EDTA, and 25 mM Tris-HCl pH 8.0.
4. Dissolve 5 mg of lysozyme in 200 µl of the buffer from step 3 and add to the mixture.
5. Place on ice for 30 min.

6. Add 2.5 ml of freshly prepared 0.2 N NaOH and 1% SDS. Place on ice for 10 min.
7. Add 1.9 ml of 3 M KAc pH 4.8. Place on ice for 15 min.
8. Centrifuge at 0°C for 30 min. at 6000 r.p.m.
9. Filter supernatant through cheese cloth.
10. Add 1 volume isopropanol and place on ice for 15 min.
11. Centrifuge at 20°C for 30 min. at 6000 r.p.m.
12. Resuspend the pellet in 300 µl of TE buffer. Transfer solution to an microcentrifuge tube. Add 300 µl LiCl and place on ice for 15 min.
13. Centrifuge at 0°C for 20 min. at 6000 r.p.m.
14. Transfer solution to a new tube and add one volume of isopropanol.
15. Place on ice for 15 min.
16. Centrifuge at 20°C for 30 min. at 13,000 r.p.m.
17. Resuspend the pellet in 150 ml of TE buffer with 3 µg RNase.
18. Incubate for 30 min. at 37°C.
19. Two times phenol:chloroform extraction.
20. One chloroform extraction.
21. Add 1/10 volume 3 M NaAcetate.

22. Precipitate DNA in two volumes of EtOH at room temperature for 10 min.
23. Centrifuge at room temperature for 20 min. at 13,000 r.p.m.
24. Resuspend the pellet in 200 µl TE buffer.

3.4.4. Subcloning DNA fragments into plasmid vectors

The following protocol was derived from Maniatis et al. (1984).

1. Mix together 200 ng of linearized plasmid DNA (phosphatase treated, Section 3.5.5) and a three-fold molar excess of the desired fragment (Section 3.3.3.) to be subcloned, total volume 8 µl.
2. Add 1 µl 10X ligase buffer (0.66 M Tris-HCl pH 7.5, 50 mM MgCl2, 50 mM dithiothreitol, 10 mM ATP) and 4 units of T4 DNA ligase. Incubate for 2 hours at 29°C for cohesive ends or 12 hours at 16°C for blunt ends.
3. Add a 2.5 µl aliquot to 100 µl of competent bacteria (Section 3.4.5). Place on ice for 10 min.
4. Heat shock the bacteria for 2 min. at 42°C.
5. Add 2 ml of L-medium and place the bacteria in a 37°C water bath for 60 min.

6. Streak 100 µl of mixture onto X-gal plates.
7. Incubate plates overnight at 37°C.
8. White and blue colonies will appear. White colonies will contain recombinant plasmids.

3.4.5. Competent bacteria

The following protocol was derived from Maniatis et al. (1984).

1. Select a single bacterial colony (Mc1022) and inoculate 5 ml of L-medium. Grow the culture overnight at 37°C.
2. Use 1 ml of the culture to inoculate 50 ml of L-medium. Grow the culture in a shaking incubator at 37°C. Grow culture to an OD650 = 0.3.
3. Chill the culture for 10 min. on ice.
4. Centrifuge the culture at 6000 r.p.m. for 5 min.
5. Resuspend the bacteria in 20 ml 0.1 M CaCl2 (ice cold) and place on ice for 30 min.
6. Repeat step 4.
7. Resuspend the bacteria in 5 ml of 0.1 M CaCl2 (ice cold) and place on ice for 24 hours.

L-medium (1 liter)
10 g peptone
5 g yeast extract
10 g NaCl
2.4 g MgSO4

X-gal plates
1 liter of L-medium
15 grams of Bactoagar
Autoclaved and cooled enough to touch then add:
150 mg ampicillin
0.04 mg of X-gal dissolved in 2 ml N,N-dimethyltomamide
0.025 mg IPTG dissolved in 2 ml H2O

3.5. Labelling DNA Fragments

3.5.1. Klenow fill in reaction

This reaction can be used to convert 5' protruding DNA to blunt ended DNA and/or to radioactive label the DNA (Maniatis et al., 1984).

1. Add 5 µl 10X of nick translation buffer (Section 3.5.3), 2 µl Klenow (2 units/µl), 5-10 µg DNA, 3 µl [α-32P]-dCTP and 5 µl 500 µM cold dNTPs if necessary. Bring total volume to 50 µl with H2O.

2. Heat for 10 min. at 37°C.
3. Stop reaction with 5 µl 0.2 M EDTA.
4. Add 50 µl 4M ammonium acetate and 140 µl EtOH.
5. Store at room temperature for 30 min.
6. Centrifuge at 13,000 r.p.m. for 30 min. at room temperature.

3.5.2. Labelling 3' recessed ends with MNLV-reverse transcriptase

The following protocol was provided by Dr. Hans Sommer (Max-Planck-Institut für Züchtungsforschung, Köln, Federal Republic of Germany).

1. Add together 10 µl (1-2 µg) DNA, 6 µl 5X buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl, 15 mM MgCl2), 1.5 µl 0.2 M DTT, 1.5 µl BSA (2 mg/ml), 3 µl [α-32P]-dATP, 7.5 µl H2O and 0.5 µl MNLV-RT (100 units/µl).
2. Incubate for 30 min. at 37°C.
3. Chase with 1 µl 10 mM dNTP and 0.5 µl MNLV-RT.
4. Incubate for 30 min. at 37°C.
5. Stop the reaction with 3 µl 0.2 M EDTA.
6. Add 35 µl 4M ammonium acetate and 140 µl EtOH.
7. Store at room temperature for 30 min.
8. Centrifuge at 13,000 r.p.m. for 30 min. at room temperature.

3.5.3. Nick translation

The following protocol was derived from Maniatis et al. (1984).

1. Add together 3 µl of 10X nick translation buffer (50 mM Tris pH 7.6, 50 mM MgCl2 and 100 mM β-Mercaptoethanol), 5 µl dNTPs (1 mM dGTP, dATP, and dTTP), 3 µl [α-32P]-dCTP, 1 µl Kornberg enzyme (2 units/µl), and 500 ng of DNA. Bring the total volume to 30 µl using H2O.
2. Place in a 16°C water bath for 2 hours.
3. Separate unincorporated [α-32P]-dCTP from the labelled fragment by a sephadex G100 column.
4. Boil the fragment for 15 min. before using for hybridization.

3.5.4. Oligo nucleotide primed labelling reaction

1. Add together 50 ng DNA, 4 µl 10X buffer [100 µl 1 mM MgCl2, 500 mM 1 M Tris-HCl (pH 7.5), 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 70 mM β-ME, 2.5 mg pd(N6)], 2 µl BSA (2 mg/ml), 3 µl [α-32P]-dATP, and 2 units of Klenow polymerase. Bring the total volume to 40 µl with H2O.
2. Incubate at room temperature for 2-4 hours.
3. Separate unincorporated [α-32P]-dCTP from the labelled fragment by a sephadex column.

4. Boil fragment for 15 min. before using for hybridization.

3.5.5. Labelling DNA with Polynucleotide Kinase

The following protocol was derived from Maniatis et al. (1984).

1. Use 10 µg of DNA fragment (Section 3.3.3.) and add 5 µl 10X CIP buffer (0.5 M Tris-HCl pH 8.0, 1 mM EDTA), 0.5 units calf intestinal alkaline phosphatase (CIP), and bring to a total volume of 50 µl with H2O. This removes 5' phosphates.
2. Incubate for 30 min. at 37°C.
3. Add 40 µl H2O, 10 µl 10X CIP stop buffer (100 mM Tris-HCl pH 8.0, 1 M NaCl, 10 mM EDTA, 100 mM EGTA) and 5 µl SDS.
4. Incubate for 45 min. at 65°C.
5. Purify DNA once with phenol:chloroform and once with chloroform.
6. Precipitate DNA with 2 volumes of EtOH, 30 min. room temperature, and 30 min. centrifugation at 13,000 r.p.m. Resuspend DNA in 10 µl TE buffer.
7. Add 5 µl 10X kinase buffer (0.5 M Tris-HCl pH 7.6, 0.1 M MgCl2, 50 mM dithiothreitol, 1 mM spermidine, 1 mM EDTA), 5 µl [γ-32P]-ATP, 10

units T4 polynucleotide kinase, and H2O to total volume of 50 µl.

8. Incubate for 30 min. at 37°C.
9. Stop reaction with 3 µl 0.2 M EDTA.
10. Add 35 µl 4M ammonium acetate and 140 µl EtOH.
11. Store at room temperature for 30 min.
12. Centrifuge at 13,000 r.p.m. for 30 min. at room temperature.

3.6. Handling Phages

3.6.1. Large scale phage preparation

1. Infect 10 ml of bacterial strain CES 200 (overnight culture) with 200 µl phage minilysate by heating at 29°C for 20 min.
2. Inoculate 1 l L-medium with bacteria culture mixture. Shake overnight at 37°C.
3. Add 10 ml chloroform and shake for 15 min.
4. Transfer the culture but not the chloroform into a new flask.
5. Add 50 µl DNase (20 mg/ml) and keep at room temperature for 30 min.
6. Add 60 g NaCl and dissolve. Place on ice for 1 hour.
7. Centrifuge for 20 min. at 7000 r.p.m.

8. Pour the supernatant into a clean flask and add 100 g PEG 1000. Place on ice for 3-12 hours.
9. Centrifuge for 15 min. at 7,000 r.p.m.
10. Dissolve the pellet in 6 ml SM buffer (10 mM Tris-HCl pH 7.5, 10 mM MgSo4, 100 mM NaCl). Store overnight at 4°C.
11. Extract two times with chloroform.
12. Add 5 g CsCl and centrifuge at 15°C overnight at 33,000 r.p.m. in a SW60 rotor.
13. Remove the white band with a pasteur pipette.

3.6.2. Extraction of phage DNA

1. Mix together 1 ml of phage (step 13 Section 3.6.1.), 4 ml 10 mM Tris-HCl (pH 7.5), 100 µl SDS, and 1 mg Proteinase K.
2. Heat for 20 min. at 37°C and then 10 min. at 50°C.
3. Extract twice with phenol:chloroform and once with chloroform.
4. Precipitate the DNA with two volumes of EtOH and let it stand at room temperature for 30 min.
5. Centrifuge for 30 min. at 8,000 r.p.m. Wash DNA with 70% EtOH and resuspend with 200 µl TE buffer.

3.7. Hybridization Techniques

3.7.1. Northerns

1. Mix together 5 µg of poly(A+) mRNA with an equal volume of loading buffer (100 µl formamide, 38 µl 38% formaldehyde and 20 µl of 10X Northern buffer). Add 1/10 volume of 50% glycerol with bromophenol blue. Northern buffer: 200 mM MOPS, 50 mM NaAcetate and 10 mM EDTA.
2. Heat the sample for 5 min. at 60°C.
3. Load the sample onto a Northern gel (2.4 g agarose, 20 ml 10X Northern buffer, 140 ml H2O and 40 ml formaldehyde).
4. 1X Northern buffer solution is added so that it reaches, but does not cover, the top of the gel. A glass plate is placed over the gel to prevent diffusion of the formaldehyde.
5. Run the gel at 100 volts for 5 hours.
6. Transfer the RNA to a nylon membrane according to the manufacturer's instructions.
7. Prehybridize the filter (300 ml 20X SSPE, 5 ml 4% ficoll, 5 ml 4% PVP, 10 ml 10% SDS, 5 mg salmon sperm DNA and H2O to one liter total volume) for four hours at 60°C in a shaking water bath.

8. Hybridize the filter in 20-40 ml of hybridization solution (75 ml 20X SSPE, 2.5 ml 4% ficoll, 2.5 ml 4% PVP, 5 ml 10% SDS, 5 mg salmon sperm DNA and H2O to 500 ml total volume), which contains the radioactive probe (Sections 3.5.3 and 3.5.4).
9. Hybridize over night at 60°C in a shaking water bath.
10. Remove the hybridization solution and wash the filter for 5 min. in 2X SSPE and 1% SDS at 60°C.
11. For very specific hybridization the filter is washed a second time as in step 10.
12. Expose the filter on X-ray film for three days.

3.7.2. Southerns

The following protocol was derived from Maniatis et al. (1984).

1. Digest the DNA with the desired restriction enzyme (10 µg for genomic DNA and 1-2 µg for plasmid DNA). After digestion, run the DNA samples on an agarose gel.
2. After running, place the gel in a 1.5 M NaCl and 0.5 M NaOH bath for one hour in order to denature the DNA.
3. Neutralize the gel in a 1 M Tris-HCl (pH 8.0) and 1.5 M NaCl bath for one hour.

4. Treat the gel for one hour in a 20X SSPE bath.
5. Transfer the DNA to a nylon membrane according to the manufacturer's instructions.
6. Prehybridize the filter for one hour at 68°C in a shaking water bath.
7. Hybridize the filter in 20-40 ml of hybridization solution, which includes the radioactive probe (Sections 3.5.3 and 3.5.4).
8. Hybridize over night at 68°C in a shaking water bath.
9. Remove the hybridization solution and wash the filter for 30 min. in 2X SSPE and 1% SDS at 68°C.
10. For very specific hybridization the filter is washed a second time with 0.2X SSPE and 0.1% SDS at 68°C for 30 min.
11. Expose the filter on X-ray film for one hour to three days, depending on the strength of the signal.

3.7.3. Phage plaques and bacterial colony lifts

1. Place a round nitrocellulose filter on the surface of a petri plate that has bacterial colonies or phage plaques.

2. Lift the filter off after a couple of minutes and place on Whatman paper that has been saturated with 1.5 M NaCl and 0.5 M NaOH.
3. After 10 min. for phage plaques and 20 min. for bacterial colonies, transfer the filter to Whatman paper saturated with 1 M Tris-HCl (pH 8.0) and 1.5 M NaCl for 10 min.
4. Transfer the filter to Whatman paper saturated with 2X SSPE for 10 min.
5. Bake the filter for 30-60 min. at 80°C.
6. Prehybridize the filter for one hour at 68°C in a shaking water bath.
7. Hybridize the filter in 20-40 ml of hybridization solution, which includes the radioactive probe (Sections 3.5.3 and 3.5.4).
8. Hybridize over night at 68°C in a shaking water bath.
9. Remove the hybridization solution and wash the filter for 30 min. in 2X SSPE and 1% SDS at 68°C.
10. For very specific hybridization the filter is washed a second time with 0.2X SSPE and 0.1% SDS at 68°C for 30 min.

11. Expose the filter on X-ray film for one hour to three days, depending on the strength of the signal.

3.8. DNA Sequencing

Two different methods of DNA sequencing were used for the c1-p allele. One complete strand from position -610 to 3274 of the C1 allele (Appendix) was sequenced using the chemical degradation method of Maxam and Gilbert (1980). The second strand and one strand from position 3274 to 3924 were mainly sequenced using the same procedure, but some areas were sequenced using the dideoxy chain termination method (Sanger et al., 1977).

The fragments sequenced were subclones from the lambda clone ligated into the vectors puc9 or puc19, or deletion derivatives were made of the subclones using Exonuclease III according to the protocol of Henikoff (1987).

The c1-m1 allele was sequenced using the dideoxy termination method and using only one strand. The lamda phage clone was subcloned into puc19 and deletion derivatives were made using Exonuclease III. The deletion derivatives were used in plasmid sequencing.

3.8.1. Chemical degradation sequencing

The fragments to be sequenced were isolated from the plasmid vectors using the appropriate restriction enzyme and separated on an agarose gel, followed by purification procedures (Section 3.3.3). The enzymes used resulted in blunt ended or 5' protruding DNA. Fragments with blunt ends were labelled using [γ-32P]-ATP in a PNK reaction. In addition to this procedure, 5' protruding fragments were labelled by filling in the 3' recessed segment with the appropriate [α-32P]-dNTP or [α-35S]-dNTP using the enzymes MNLV-RT reverse transcriptase (Section 3.5.2) or Klenow polymerase (Section 3.5.1). In all instances 5 to 10 µg of DNA were labelled.

1. Dissolve the labelled DNA in 34 µl H20 and divide the solution into the appropriate base specific reaction tubes (5 µl for G and C, 8 µl for A and C-T, and 8 µl for reserve in case of mistakes).
2. To the G reaction tube add 200 µl of cacodylate buffer (50 mM sodium-cacodylate pH 8.0, 10 mM MgCl2 and 1 mM EDTA) and heat for 5 min. at 37°C. Add 1 µl 100% DMS, mix by vortexing and heat at 37°C for 25 seconds. Add 50 µl stop buffer (1.5 M NaAcetate ph 7.0 and 1 M ß-Mercaptoethanol), 750 µl cold EtOH, and mix.

3. To the A reaction tube add 100 µl start buffer (1.2 N NaOH and 1 mM EDTA) and heat for 6 min at 95°C in a rack that has a tight sealing lid to prevent the lid of the tube from opening. Add 150 µl stop buffer (1 M AcOH and 1 mM EDTA), 750 µl cold EtOH and mix.
4. To the C reaction tube add 20 µl 5 M NaCl, 30 µl hydrazine and heat for 13 min. at 37°C. Add 200 µl 0.3 M NaAc pH 6.5, 750 µl EtOH and mix.
5. To the C-T reaction tube add 20 µl H2O, 30 µl hydrazine and heat for 10 min. at 37°C. Add 200 µl 0.3 M NaAc ph 6.5, 750 µl EtOH and mix.
6. Place all tubes in a -70°C freezer for 30 min. to precipitate the DNA. Centrifuged the tubes at 13,000 r.p.m. at 4°C for 30 min. Wash the pellet with 80% EtOH and air dry.
7. Resuspend the pellet in 80 µl 1 M piperidine and shake on an Eppendorf shaker for 10 min.
8. Heat the tubes at 95°C for 30 min. in a rack that has a tight fitting top in order to prevent the lid of the tube from opening.
9. Make needle size holes in the lid of the tubes.
10. Freeze the tubes in liquid nitrogen and lyophilized in a Speedvac for 2 hours.

11. Resuspend the pellet is in 50 µl H2O and shake for 10 min. Repeat step 10.
12. Repeat step 11.
13. Measure the level of radioactivity for each tube and then add enough loading buffer so that 1 µl will contain the proper level of radioactivity (8000 cpm for G and C, and 10,000 cpm for A and C-T).
14. Heat samples for 90 seconds at 95°C.
15. Load 0.5-1.0 µl of sample on a sequencing gel.

Sequence Loading Buffer
80% deionized formamide
50 mM Tris-borate pH 8.3
0.1% Xylene cyanol
0.1% Bromophenol blue

3.8.2. Dideoxy termination sequencing

All dideoxy sequencing was done using a commercial kit obtained from Pharmacia, which contains all necessary chemical compounds except [α-35S]-dATP and uses T7 polymerase instead of Klenow for the primer extension. The procedure used was a slight modification of the instructions provided by the manufacturer. The major

change is that the time for the DNA strand synthesis is shortened.

1. One and one half µg of plasmid DNA (Section 3.4.3) is denatured using a NaOH treatment (8 µl of DNA and 2µl of 2 M NaOH) for 10 min at room temperature.
2. Precipitate the DNA by adding 3 µl 3 M sodium acetate, 7 µl H2O and 60 µl EtOH. Mix the solution and place in a -70°C freezer for 15 min. Centrifuge the tube for 10 min. at 13,000 r.p.m. Remove the supernatant and wash the pellet once with 100% EtOH and once with 70% EtOH.
3. Redissolve the DNA in 10 µl of H2O for 10 min.
4. Add 2 µl of annealing buffer and 2 µl of primer (1 µl is equal to 0.5 pica moles). Heat sample for 20 min. at 37°C and then let cool at room temperature for 10 min.
5. For each base specific reaction tube add 2.5 µl of the proper termination mixture and preheat at 37°C. This solution contains the specific dideoxy nucleotide to terminate DNA strand synthesis.
6. To the tube containing the DNA, add 3 µl labelling mixture, 1 µl [α-35S]-dATP, and 2 µl T7

polymerase (1.5 units/µl) to separate locations on the side of the tube. Start DNA strand synthesis by centrifuging the solutions together. After 2 min., take 4.5 µl aliquots from this tube and add to the side of the base specific reaction tubes. Mix by centrifugation.

7. After 5 min., add 5 µl of the stop buffer to the side of the base specific reaction tubes. Mix by centrifugation.
8. Heat the samples 90 seconds at 80°C before loading them on the sequencing gels. Load an 1.5 aliquot of the sequencing reaction to separate slots on the gel.

3.8.3. Preparation of the glass plates

In the preparation of sequencing gels the glass plates, between which the acrylamide is poured, is one of the most critical steps. In this procedure the ear plate is treated so that the acrylamide will not adhere while the back plate is prepared so the acrylamide will adhere. After running, the gels are dried onto the plate and exposed to X-ray sensitive film.

1. Wash all plates with soap and water to remove any old acrylamide and rinse well with distilled H2O.

2. Wash back and ear plate twice with chloroform and once with EtOH.
3. Lay back plate on a level surface and cover desired side with silan (25 ml EtOH, 100 µl GF31 Silan from Wacker Silcone and 750 µl 10% AcOH). Let dry for 15-30 min.
4. Remove excess silan by holding the back plate at an angle and spray at least seven times with EtOH.
5. Lay ear plate on a level surface and apply 5 ml of dichlorosilan (5% dichlorosilan in CCl4) a total of three times across the face of plate using a lint fee tissue.
6. Remove excess silan from the ear plate by rubbing the surface hard with a lint free tissue saturated with EtOH. Repeat for a total of three times.
7. Air dry the plates and inspect for lint before assembly. Assemble the plates by placing the treated sides face to face and separate them with a 0.2 mm thick spacer on the two edges of the plate. Clamp the gels together and seal the bottom with tape.

8. Tilt the plates at an angle and slowly pour the acrylamide mixture down one of the edges, until the space between the plates is filled. Lay the plates on a level surface and place a slot former in the top. Allow the acrylamide to polymerize.

3.8.4. Acrylamide mixtures for sequencing

Different percentages of acrylamide gels and the length of run are used to read the maximum number of bases. High percentage acrylamide gels (16% and 8%) and short runs are used to read the first 200 bp of a fragment, while long runs of low percentage gels (6% and 5%) are used to read bases from 180 to 400.

A basic stock of 40% acrylamide is used in all gels. The stock is prepared by adding 380 g of 2X acrylamide, 20 g of bis-acrylamide and two teaspoons of amberlite to one liter of H2O. The mixture is slowly stirred overnight at 4°C in the dark. The solution is filtered twice. Once through a normal paper filter and once using a Whatman glass microfiber filter under low vacuum pressure.

10X TBE Buffer Stock Solutions (1 liter)

Chemical	acrylamide gels ≤ 6%	acrylamide gels ≥ 8%
Tris	162.0 g	60.5 g
Boric Acid	27.5 g	30.9 g
EDTA	9.5 g	3.72 g

Acrylamide gel mixtures

gel %	urea	40% acrylamide	10X TBE	1% APS
16%	24 g	20.0 ml	7.5 ml	2.5 ml
8%	24 g	10.0 ml	7.5 ml	2.5 ml
6%	24 g	7.5 ml	5.0 ml	2.5 ml
5%	24 g	6.3 ml	5.0 ml	2.5 ml

Filter all solutions and vacuum the solution for an additional 30 seconds after is starts to boil. Add 27 µl β-ME, mix quickly and pour gel.

3.8.5. Running sequence gels

There are several sequencing gel combinations that can be used to achieve the same desired results. Mainly, the procedure used depends on a particular situation. The major way is to run a sequencing reaction on several gels (40 cm in length) of different percentages for distinct

periods of time. The time of the run is determined by the migration of the dye markers xylene cyanol (green) and bromophenol blue.

For 32P chemical degradation reactions, four different percentages of gels are used. A 16% gel is used for the first 45 bp (blue dye 12 cm from the bottom) 8% for bp 35-200 (green dye 15 cm from bottom) 6% for bp 175-325 (green dye 15 cm from the top after the second loading of the dye) and 5% for bp 275-400 (green dye run out of the bottom after the second loading of the dye).

In contrast, only two gels are used for the for the dideoxy chain termination (35S) method; an 8% gel for bases 1-200 (blue dye 5 cm from the bottom) and 6% gel for bases 175-350 (a second loading of dye markers and the green dye 15 cm from the top). The reason that fewer gels are needed is that when 35S is used, the sequencing bands labeled with 35S are less diffused than those labeled with 32P. This allows for the reading of more bases at the top of the gel, where the spacing between bases is very small.

After the acrylamide has polymerized the slot former is removed and the slots are cleaned by spraying H2O into the slots with a syringe. The gel is placed in the gel tank and the tank is filled with the appropriate running buffer. For chemical degradation reactions the slots are

filled with 8 M urea dissolved in the same running buffer. The gels are preheated to 55°C. Before loading the sequencing reactions the slots are cleaned again with a syringe filled with the running buffer. All gels are ran at the maximum voltage that the power supply will provide (5000 or 3000 volts) and with enough current to maintain the gel temperature at 50-55°C.

After running, the gel is removed from the tank and cool water is poured over the ear plate. The spacer is then removed and a razor blade is placed at one of the bottom corners between the two plates and twisted in order to separate the ear plate from the acrylamide gel. The back plate with the attached acrylamide gel is placed in a container that has 10% acetic acid and shaken. After 60 min. the plate is dried vertically in an 80°C oven for one hour. The plate is allowed to cool and then exposed on X-ray film. All gels except the 16% are treated in this manner. Instead of the acetic acid and drying treatment, the back plate with the 16% gel are covered with plastic wrap, exposed to X-ray film and placed in the -70°C freezer.

3.8.6. Construction of deletion derivatives using Exonuclease III

The method described by Henikoff (1987) was used to make plasmid deletion derivatives for sequencing. The general concept is cut the plasmid containing the insert with a restriction enzyme that leaves 5' protruding or blunt ends. The DNA is then digested with Exonuclease III. This method takes advantage of the properties of Exonuclease III which digests double stranded DNA from the 3' end at a uniform rate and the rate of digestion is regulated by the temperature. Unidirectional digestion can be achieved by blocking one end of the double stranded DNA with a 3' protruding end or filling in a 3' recessed end with α-phosphorothioate deoxynucleoside triphosphates.

1. Dissolve 5-10 µg DNA with 60 µl Exo buffer (66 mM Tris-HCl, 0.66 mM MgCl2).
2. Heat DNA to desired temperature (34-37°C).
3. Add 500 units to the DNA solution and quickly mix.
4. Every 30 seconds, remove a 2.5 µl aliquot and place it into a prepared tube containing 7.5 µl S1 nuclease buffer (172 µl H2O, 27 µl 10X S1 buffer, 60 units S1 nuclease). 10X S1 buffer:

1.1 ml 3 M KOAc (pH 4.6), 5 ml 5 M NaCl, 5 ml glycerol, 30 mg ZnSO4.

5. After desired number of aliquots are taken, remove the samples from the ice and place at room temperature for 30 min.
6. Add 1 µl S1 stop buffer (0.3 M Tris Base, 0.05 M EDTA) and place samples in 70°C water bath for 10 min.
7. Transfer aliquots to a 37°C water bath.
8. Add 1 µl Klenow mix (3 µl 0.1 M Tris-HCl pH 8, 6 µl 1 M MgCl2, 20 µl H2O, 3 units of Klenow polymerase) and heat for 2-5 min.
9. Add 1 µl 0.125 mM dNTPs and heat for 5 min.
10. Place the samples at room temperature for 30 min. and add 40 µl ligase mix (0.8 ml H2O, 0.2 ml 5X ligase buffer, 5 units T4 DNA ligase). 5X ligase buffer: 2.5 ml 1 M Tris-HCl (pH 7.6), 0.25 ml 1 M MgCl2, 0.5 ml 0.1 M dithiothreitol. 0.5 ml 0.1 M ATP, 2.5 polyethylene glycol 6000, water to 10 ml.
11. Take a 10 µl aliquot and add to 200 µl competent bacteria (Section 3.4.5). Transform bacteria as in Section 3.4.4.

4. RESULTS

4.1. Molecular Analysis of the c1-p Allele

In order to understand the mechanism conferring the receptivity to light induction of the c1-p allele, it was necessary to determine the DNA sequence of c1-p. To pursue this objective, the c1-p allele was cloned in our laboratory by Dr. Javier Paz-Ares. He used a homozygous c1-p plant as the DNA source to construct a MboI partial genomic library in the lambda phage EMBL4. A phage carrying the c1-p allele was identified by using probes that were specific for the C1 locus.

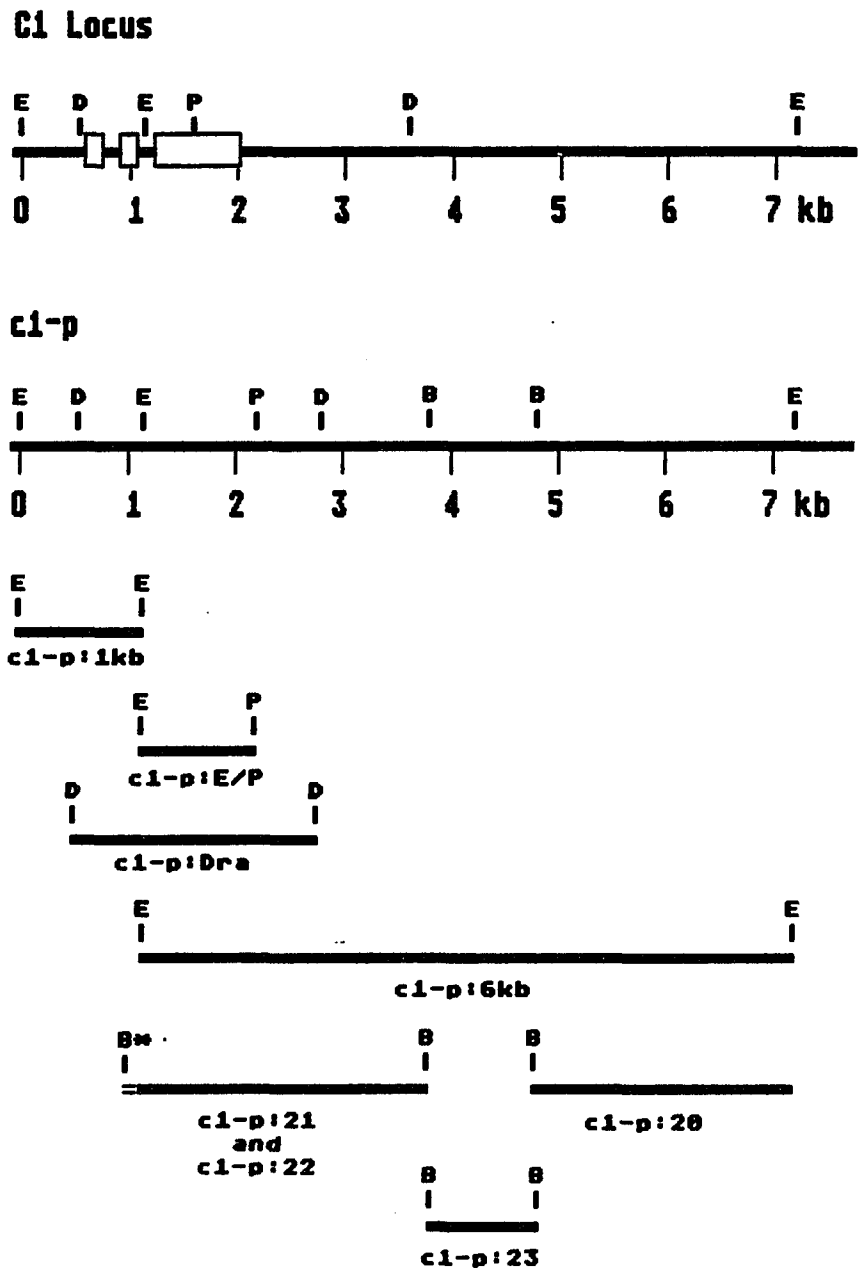
The phage clone was given to the author for molecular analysis. The DNA structure of the c1-p allele was determined by subcloning portions of the phage clone into plasmid vectors and subsequent sequence analysis.

4.1.1. Construction of subclones using plasmid vectors

The restriction map of the wild type C1 allele (Figure 4.1) was used as a guide to determine which restriction enzymes could be used to make subclones of the lambda phage clone containing the c1-p allele.

The lambda DNA was purified, as described in the material and methods (Section 3.6), and digested with the desired restriction enzymes according to the

Figure 4.1. Restriction map of C1 and c1-p. The restriction sites are B = BamHI, D = DraI, E = EcoRI, and P = PstI. The open boxes represent the three exons of C1. Subclones of the c1-p genomic phage clone are listed below the restriction map. Subclones c1-p20 to c1-p23 were derived from c1-p6kb. After digesting c1-p6kb with BamHI, a fragment was still attached to puc19, and this construct was religated to give c1-p20. "B*" indicates the BamHI site in the puc9 polylinker was used


C1 Locus
E D E P D E
0 1 2 3 4 5 6 7 kb
c1-p
E D E P D B B E
0 1 2 3 4 5 6 7 kb
E E
c1-p:1kb
E P
c1-p:E/P
D D
c1-p:Dra
E E
c1-p:6kb
B* B B
c1-p:21
and
c1-p:22
c1-p:20
B B
c1-p:23

manufacturers' instructions. The DNA fragments were separated on an agarose gel, isolated and purified (Section 3.3). Subclones were made by ligating the fragments into the plasmid vectors pUC9 or pUC19 (Vieira and Messing, 1982) (Section 3.4.4).

Four subclones of the lambda clone were originally constructed (Figure 4.1). The first two were made using the restriction enzyme EcoRI, which yielded a 1.1 and 6 kb fragments. These were ligated into the EcoRI site of pUC9. It was anticipated that the 1.1 kb fragment would correspond to the 1.1 kb (c1-p:1kb) fragment of the wild type allele, which ranges from the promoter to the middle of intron 2, and the 6 kb (c1-p:6kb) fragment would cover the remaining 3' end including flanking DNA. c1-p:6kb was further subcloned by cutting this plasmid with BamHI and subcloning two of the fragments into the BamHI site of pUC19 (c1-p:21, c1-p:22 and c1-p:23). The remaining fragment was still attached to pUC9 and religated to give a new subclone (c1-p:20). The third subclone (c1-p:E/P) was made by a double digestion of the phage clone with EcoRI and PstI into the EcoRI/PstI sites of pUC9. In theory, this digestion would yield a 500 bp fragment that would cover part of intron two and most of the third intron. A 1 kb fragment was isolated instead of the

anticipated 500 bp fragment, indicating that there was an insert in the c1-p allele or that some other sequence alterations had occurred. The fourth subclone (c1-p:Dra) was constructed by using DraI and ligating the fragment into the SmaI site of pUC19. This subclone was made so that there would be an overlapping clone with the EcoRI subclones for sequence analysis. The DraI fragment was about 400 bp smaller than expected. Figure 4.1 gives a schematic diagram of the restriction map for the c1-p and wild type C1 alleles.

4.1.2. Sequencing of the c1-p allele

Two different methods of DNA sequencing were used for the c1-p allele. One strand corresponding to positions -610 to 3274 of the C1 allele (Appendix) was sequenced using the chemical degradation method of Maxam and Gilbert (1980) (Section 3.8.1). The second strand and one strand from position 3274 to 3924 were mainly sequenced using the same procedure, but some areas were sequenced using the dideoxy chain termination method (Sanger et al., 1977) (Section 3.8.2).

The fragments that were sequenced were derived from the subclones described in Section 4.1.1, or deletion derivatives were made of the subclones using Exonuclease

III according to the protocol of Henikoff (1987) (Section 3.8.6).

4.1.2.1. Construction of deletion derivatives using Exonuclease III The method described by Henikoff (1987), Section 3.8.6, was used to make plasmid deletion derivatives for sequencing. This method takes advantage of the properties of Exonuclease III which digests double stranded DNA from the 3' end at a uniform rate and the rate of digestion is regulated by the temperature.

Unidirectional digestion can be achieved by blocking one end of the double stranded DNA. This can be done by cutting the DNA with a restriction enzyme that leaves a 4 bp 3' protruding end (Henikoff, 1987) or filling in a 5' protruding end with α-phosphorothioate deoxynucleoside triphosphates (Putney et al., 1981).

In the case of plasmid c1-p:1kb the DNA was digested first with the restriction enzyme PstI and then with BamHI. Plasmids c1-p:21 and c1-p:22 (BamHI subclones of c1-p6kb in opposite orientations) were digested with SphI. c1-p:22 was then digested with XhoI while plasmid c1-p:21 was digested with XbaI. The fragments were separated on an agarose gel and the desired fragment was electroeluted (Section 3.3; Maniatis et al., 1984). Following purification steps (Section 3.3), all desired DNA

fragments from plasmids c1-p:1kb, c1-p:21 and c1-p:22 were digested using the Exonuclease III digestion protocol (Section 3.8.6).

4.1.2.2. Chemical degradation sequencing The chemical degradation sequencing method of Maxam and Gilbert (1980) was the primary method used to sequence the c1-p allele. Because deletions can occur in M13 vectors used in the dideoxy termination method, it was decided that the chemical degradation sequencing method would be the most reliable technique.

The fragments to be sequenced were isolated from the plasmids c1-p:1kb, c1-p:Dra, c1-p:E/P, c1-p:21, c1-p:22 and the deletion derivative clones (Section 4.1.2.1) using the appropriate restriction enzymes. The fragments were isolated and purified (Section 3.3) then radioactive labelled (Sections 3.5 and 3.8.1).

When the DNA could not be specifically labelled at one end of the DNA molecule, such as in the case of PNK labelling, the fragment was digested again with a restriction enzyme that cut internally. The fragments were separated on an agarose gel, electroeluted, and purified using a RPC5 column (Section 3.3).

4.1.2.3. Dideoxy termination sequencing Dideoxy termination sequencing was adopted in the later stages of

the sequence analysis of the c1-p allele, because of the commercial availability of T7 polymerase. This enzyme is much more efficient and dependable than the Klenow fragment of E. coli DNA polymerase that is normally used.

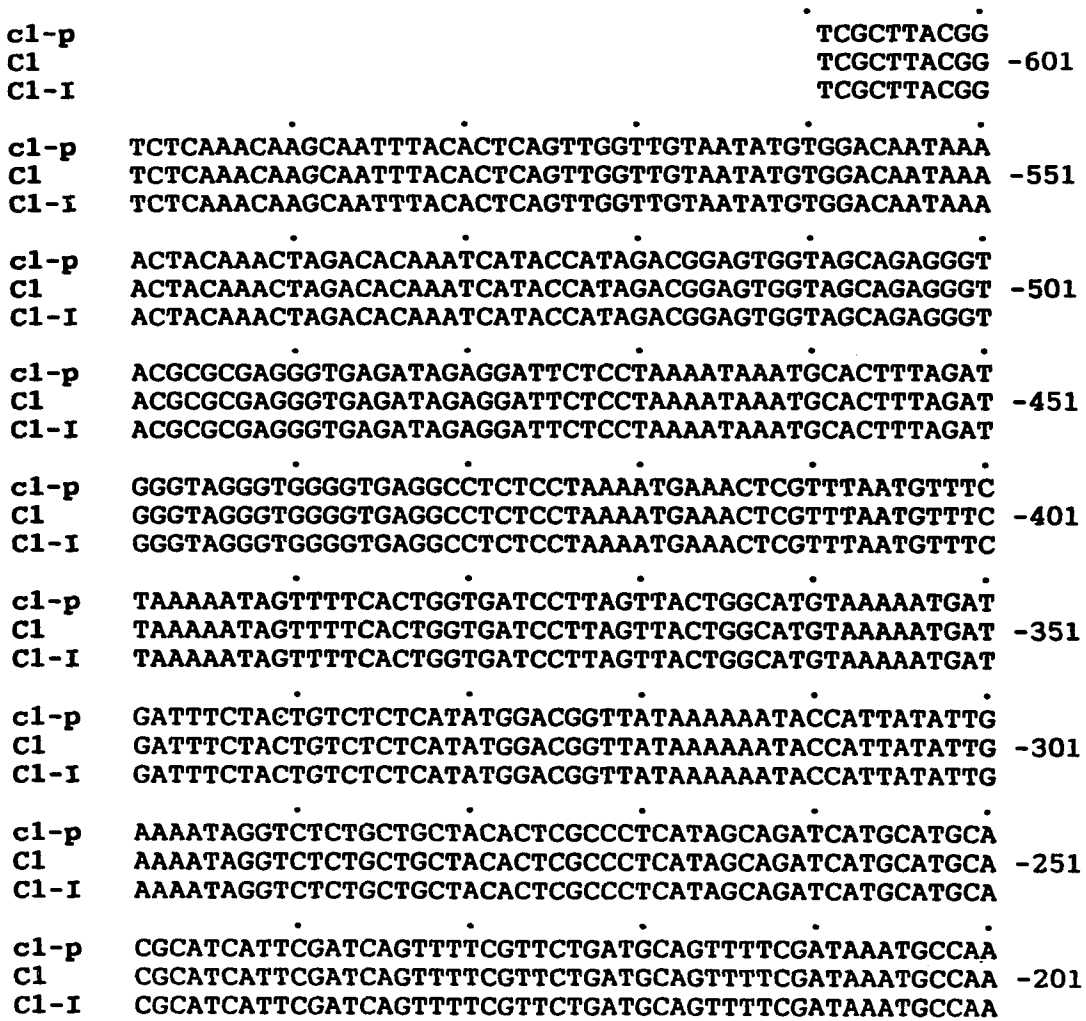
By using T7 polymerase, it was possible to do all termination sequencing on plasmid vectors and not single stranded vectors. In this way, deletions could be avoided that commonly occur in single stranded vectors. The majority of the chemicals were provided in a kit obtained from Pharmacia. The manufacturer's protocol, with minor modifications (Section 3.8.2), was used.

4.1.3. Sequence comparisons of c1-p to C1 and C1-I

The DNA sequence comparisons of the promoter, introns and exons of c1-p, C1 and C1-I are shown in Figure 4.2. A comparison of the 3' regions of C1 and c1-p are shown in Figure 4.3. No comparisons could be made with C1-I in the 3' region because there is a large insert in the C1-I allele 8 bp from the translation stop site and there was no sequence data available on the other side of the insert.

In contrast to the rest of the sequence, the promoter and the first two exons of all alleles are highly conserved (Figure 4.2). In the promoter (-610 to -1 in C1), there are only two locations (-123 and -152) where

Figure 4.2. A comparison of the DNA sequences of C1, C1-I and c1-p from -610 to 1086 of the C1 locus. The start and stop sites of translation, TATA box, and CAAT box are boxed in. Introns are presented using lower case letters. Differences between the C1 versus c1-p or C1-I are highlighted in black. An asterisk indicates the sequence of the three alleles is different at that location


c1-p TCGCTTACGG
C1 TCGCTTACGG -601
C1-I TCGCTTACGG
c1-p TCTCAAACAAGCAATTTACACTCAGTTGGTTGTAATATGTGGACAATAAA
C1 TCTCAAACAAGCAATTTACACTCAGTTGGTTGTAATATGTGGACAATAAA -551
C1-I TCTCAAACAAGCAATTTACACTCAGTTGGTTGTAATATGTGGACAATAAA
c1-p ACTACAAACTAGACACAAATCATACCATAGACGGAGTGGTAGCAGAGGGT
C1 ACTACAAACTAGACACAAATCATACCATAGACGGAGTGGTAGCAGAGGGT -501
C1-I ACTACAAACTAGACACAAATCATACCATAGACGGAGTGGTAGCAGAGGGT
c1-p ACGCGCGAGGGTGAGATAGAGGATTCTCCTAAAATAAATGCACTTTAGAT
C1 ACGCGCGAGGGTGAGATAGAGGATTCTCCTAAAATAAATGCACTTTAGAT -451
C1-I ACGCGCGAGGGTGAGATAGAGGATTCTCCTAAAATAAATGCACTTTAGAT
c1-p GGGTAGGGTGGGGTGAGGCCTCTCCTAAAATGAAACTCGTTTAATGTTTC
C1 GGGTAGGGTGGGGTGAGGCCTCTCCTAAAATGAAACTCGTTTAATGTTTC -401
C1-I GGGTAGGGTGGGGTGAGGCCTCTCCTAAAATGAAACTCGTTTAATGTTTC
c1-p TAAAAATAGTTTTCACTGGTGATCCTTAGTTACTGGCATGTAAAAATGAT
C1 TAAAAATAGTTTTCACTGGTGATCCTTAGTTACTGGCATGTAAAAATGAT -351
C1-I TAAAAATAGTTTTCACTGGTGATCCTTAGTTACTGGCATGTAAAAATGAT
c1-p GATTTCTACTGTCTCTCATATGGACGGTTATAAAAAATACCATTATATTG
C1 GATTTCTACTGTCTCTCATATGGACGGTTATAAAAAATACCATTATATTG -301
C1-I GATTTCTACTGTCTCTCATATGGACGGTTATAAAAAATACCATTATATTG
c1-p AAAATAGGTCTCTGCTGCTACACTCGCCCTCATAGCAGATCATGCATGCA
C1 AAAATAGGTCTCTGCTGCTACACTCGCCCTCATAGCAGATCATGCATGCA -251
C1-I AAAATAGGTCTCTGCTGCTACACTCGCCCTCATAGCAGATCATGCATGCA
c1-p CGCATCATTCGATCAGTTTTCGTTCTGATGCAGTTTTCGATAAATGCCAA
C1 CGCATCATTCGATCAGTTTTCGTTCTGATGCAGTTTTCGATAAATGCCAA -201
C1-I CGCATCATTCGATCAGTTTTCGTTCTGATGCAGTTTTCGATAAATGCCAA

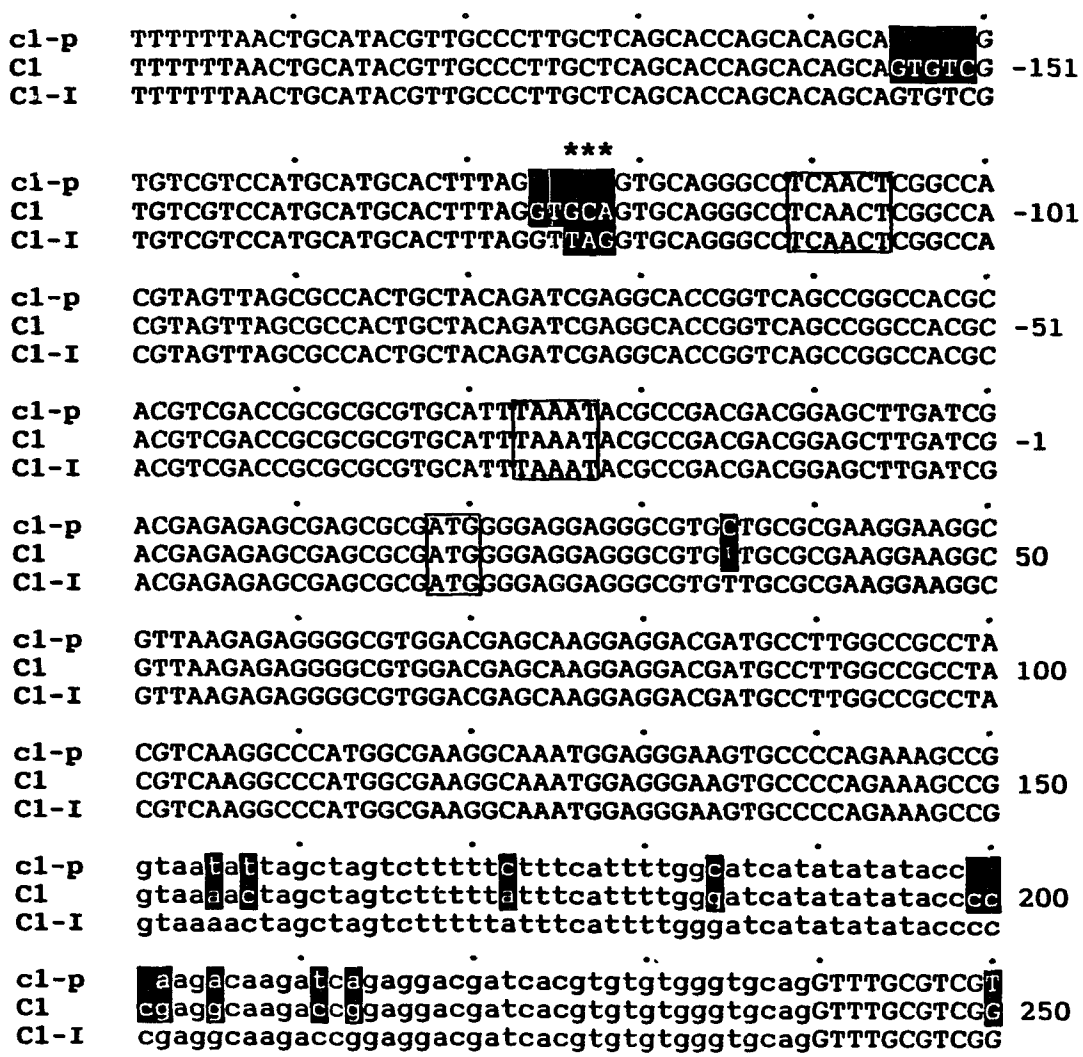


Figure 4.2. Continued

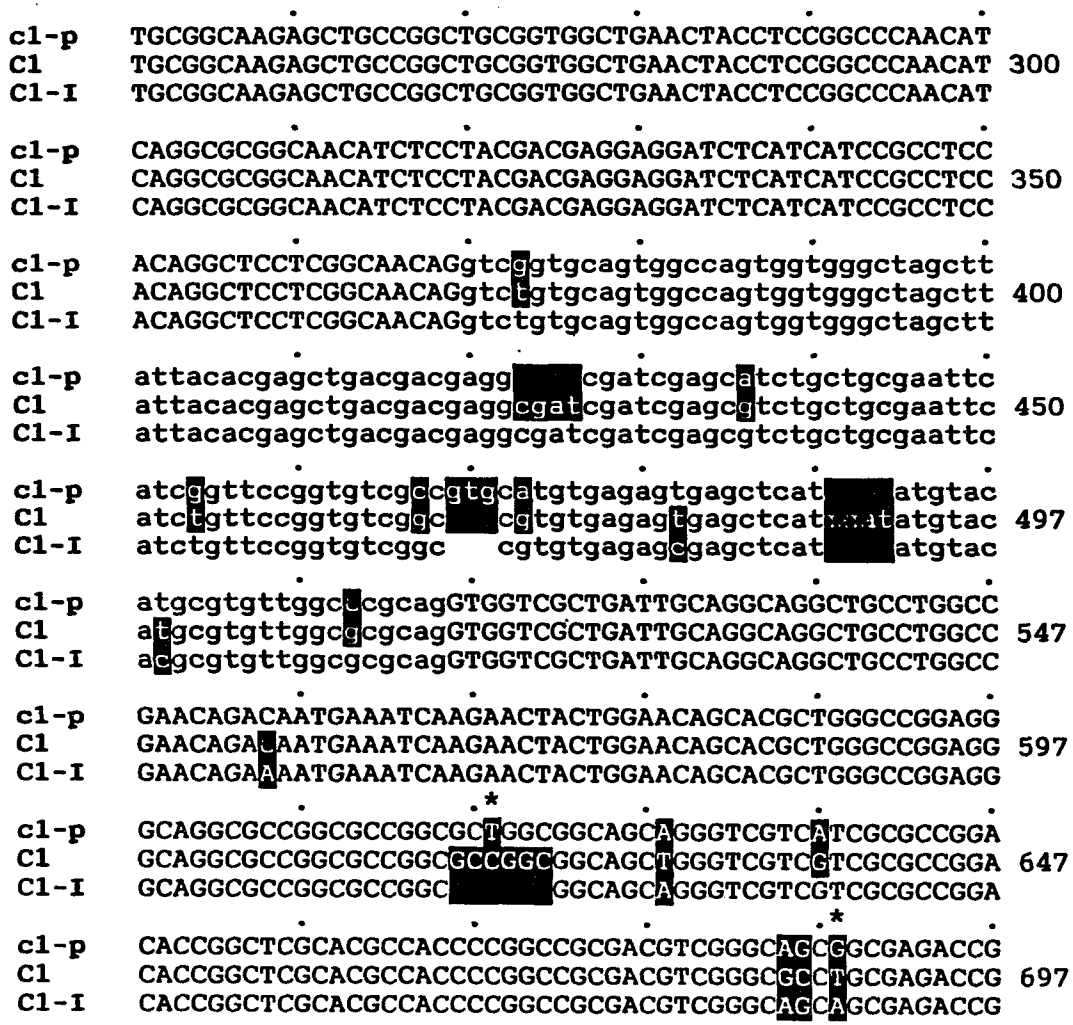


Figure 4.2. Continued

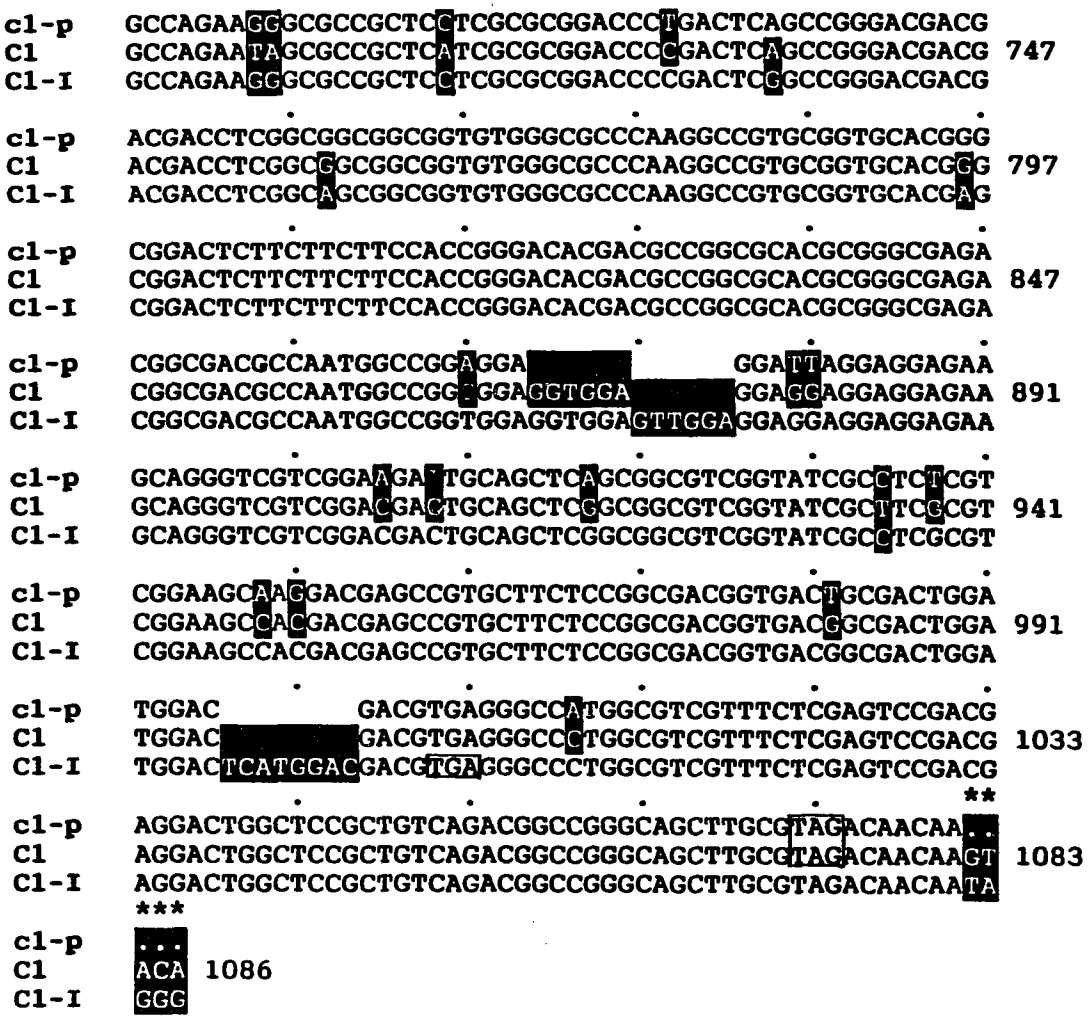


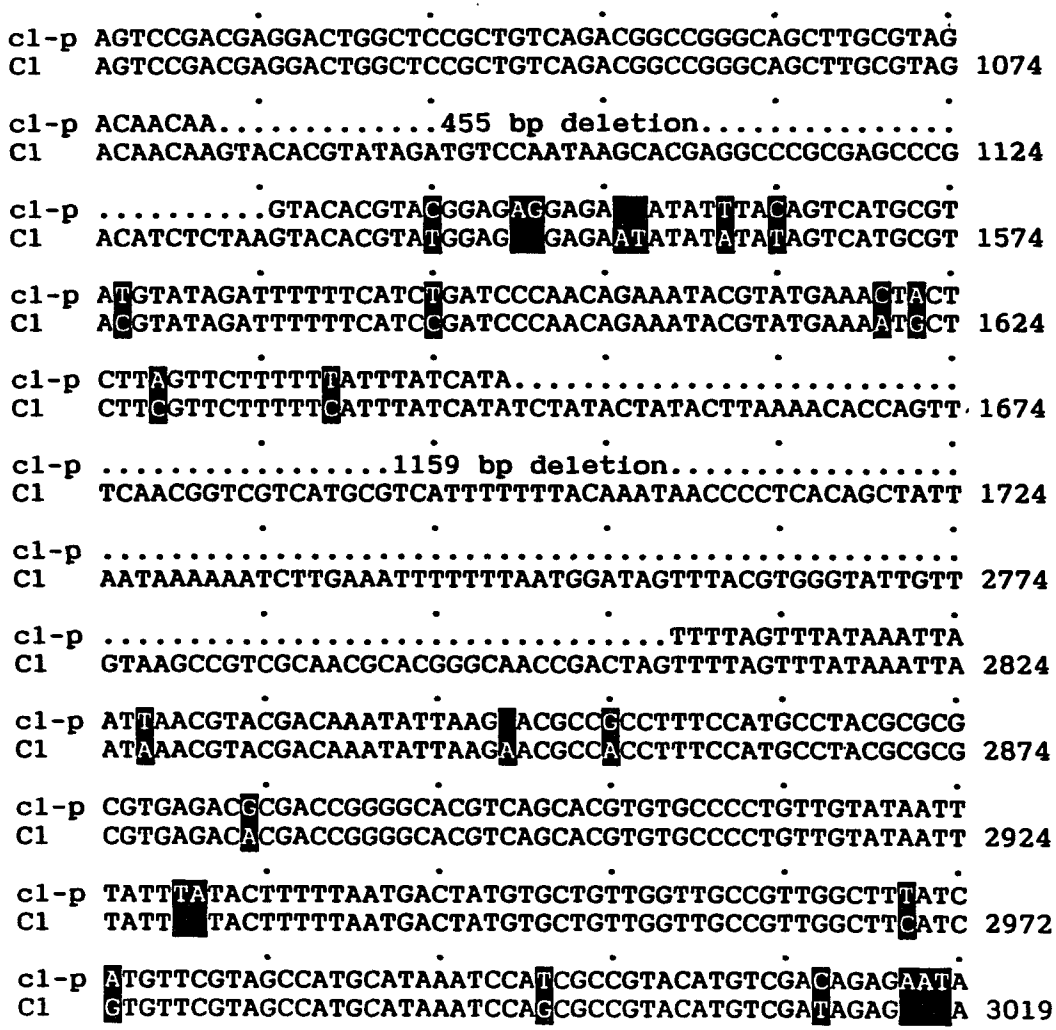
Figure 4.2. Continued

changes have occurred and all of these are in partially repetitive segments. These alterations may be due to slippage during DNA replication or the insertion and excision of a transposable element. Similar changes are found in other parts of the c1-p allele (at positions 199, 488, and 872 of C1).

There are a number of base pair alterations in both introns and in the third exon of c1-p and C1-I when compared with the C1 allele as a standard (Paz-Ares et al., 1987). In total, the c1-p allele has more alterations than the C1-I allele, but both alleles share several alterations (positions 488, 628, 685, 686, 705, 706, 716, and 935 in C1) or have changes in the same area (positions -127, 619, 688, 871, and 1091 in C1). These results indicated that c1-p and C1-I are more closely related to each other than to C1 and that some sequences may be more susceptible to DNA modification.

There are several base pair differences between the c1-p and the C1 allele in the 3' region after the translation stop site (Figure 4.3). The most important alterations are three large deletions in the c1-p allele (Figures 4.3 and 4.4). One deletion (DEL1) is 455 bp (position 1082 to 1536 in the C1 allele). In the C1 allele, DEL1 covers the area where the poly(A) addition

Figure 4.3. A comparison of the 3' DNA sequence of the wild type C1 and c1-p alleles from 1025 to 3924 of the C1 locus. Minor base pair differences are highlighted in black. Large deletions in the c1-p sequence are marked with dots and the size of the deletion is given


c1-p AGTCCGACGAGGACTGGCTCCGCTGTCAGACGGCCGGGCAGCTTGCGTAG
C1 AGTCCGACGAGGACTGGCTCCGCTGTCAGACGGCCGGGCAGCTTGCGTAG 1074
c1-p ACAACAA.....455 bp deletion.....
C1 ACAACAAGTACACGTATAGATGTCCAATAAGCACGAGGCCCGCGAGCCCG 1124
c1-p .....GTACACGTACGGAGAGGAGA ATATTTACAGTCATGCGT
C1 ACATCTCTAAGTACACGTATGGAG GAGAATATATATATAGTCATGCGT 1574
c1-p ATGTATAGATTTTTTCATCTGATCCCAACAGAAATACGTATGAAACTACT
C1 ACGTATAGATTTTTTCATCCGATCCCAACAGAAATACGTATGAAAATGCT 1624
c1-p CTTAGTTCTTTTTTATTTATCATA.....
C1 CTTCGTTCTTTTTTCATTTATCATATCTATACTATACTTAAAACACCAGTT 1674
c1-p .....1159 bp deletion.....
C1 TCAACGGTCGTCATGCGTCATTTTTTTACAAATAACCCCTCACAGCTATT 1724
c1-p .....
C1 AATAAAAAATCTTGAAATTTTTTTAATGGATAGTTTACGTGGGTATTGTT 2774
c1-p .....TTTTAGTTTATAAATTA
C1 GTAAGCCGTCGCAACGCACGGGCAACCGACTAGTTTTAGTTTATAAATTA 2824
c1-p ATTAACGTACGACAAATATTAAG ACGCCGCCTTTCCATGCCTACGCGCG
C1 ATAAACGTACGACAAATATTAAGAACGCCACCTTTCCATGCCTACGCGCG 2874
c1-p CGTGAGACGCGACCGGGGCACGTCAGCACGTGTGCCCCTGTTGTATAATT
C1 CGTGAGACACGACCGGGGCACGTCAGCACGTGTGCCCCTGTTGTATAATT 2924
c1-p TATTTATACTTTTTAATGACTATGTGCTGTTGGTTGCCGTTGGCTTTATC
C1 TATT TACTTTTTAATGACTATGTGCTGTTGGTTGCCGTTGGCTTCATC 2972
c1-p ATGTTCGTAGCCATGCATAAATCCATCGCCGTACATGTCGACAGAGAATA
C1 GTGTTCGTAGCCATGCATAAATCCAGCGCCGTACATGTCGATAGAG A 3019

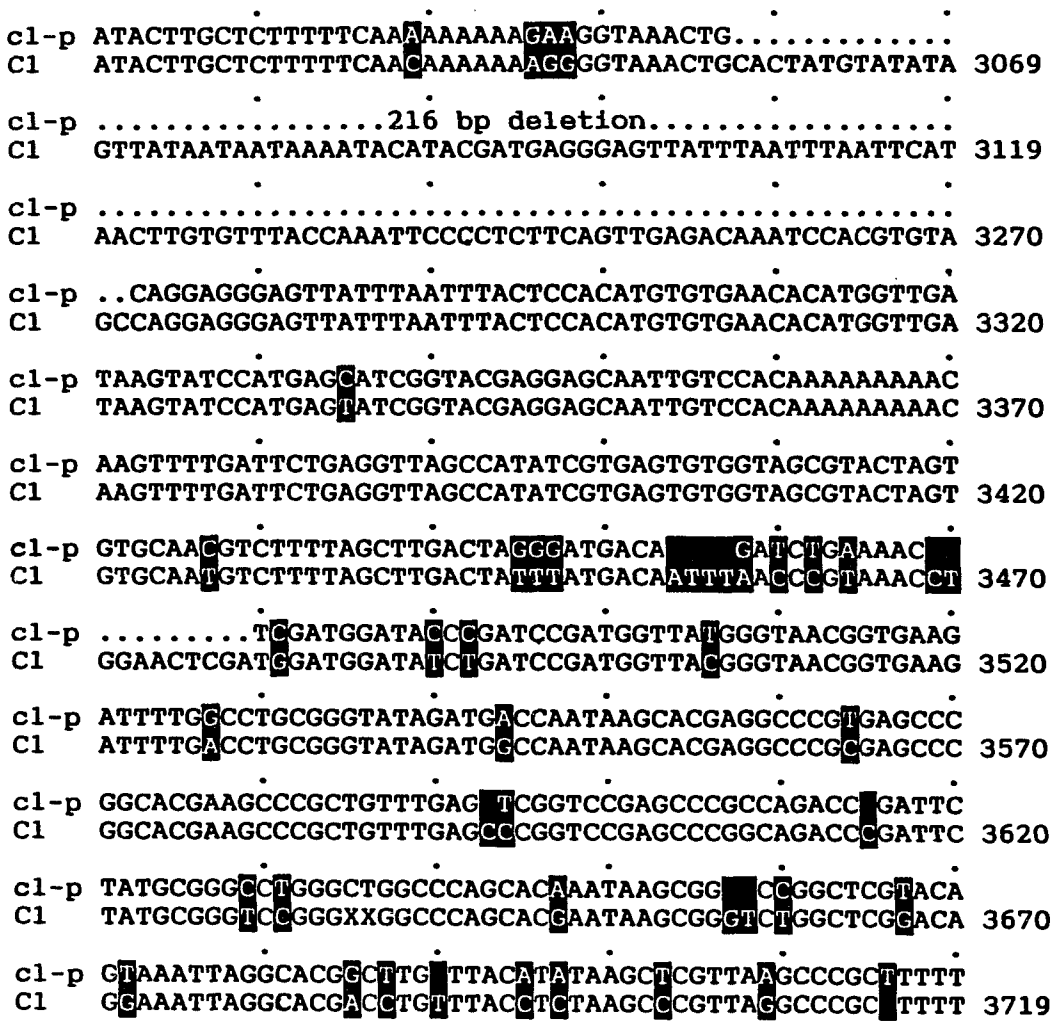


Figure 4.3. Continued

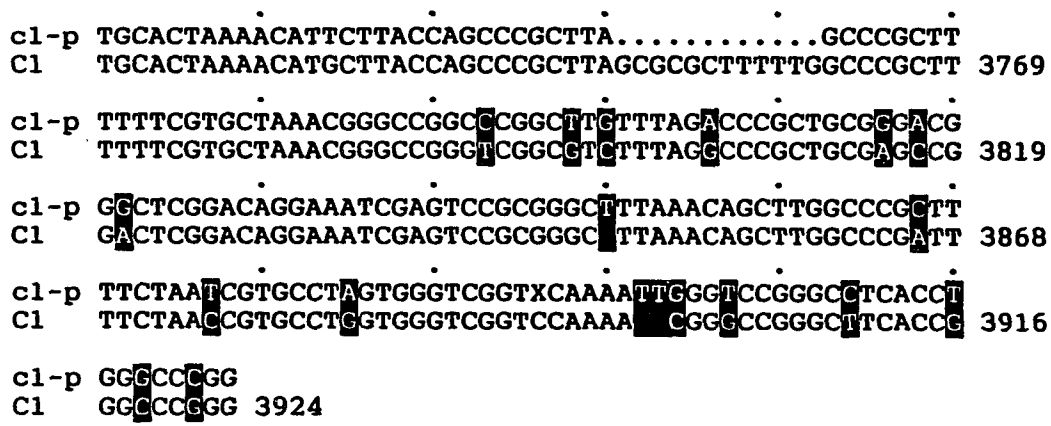
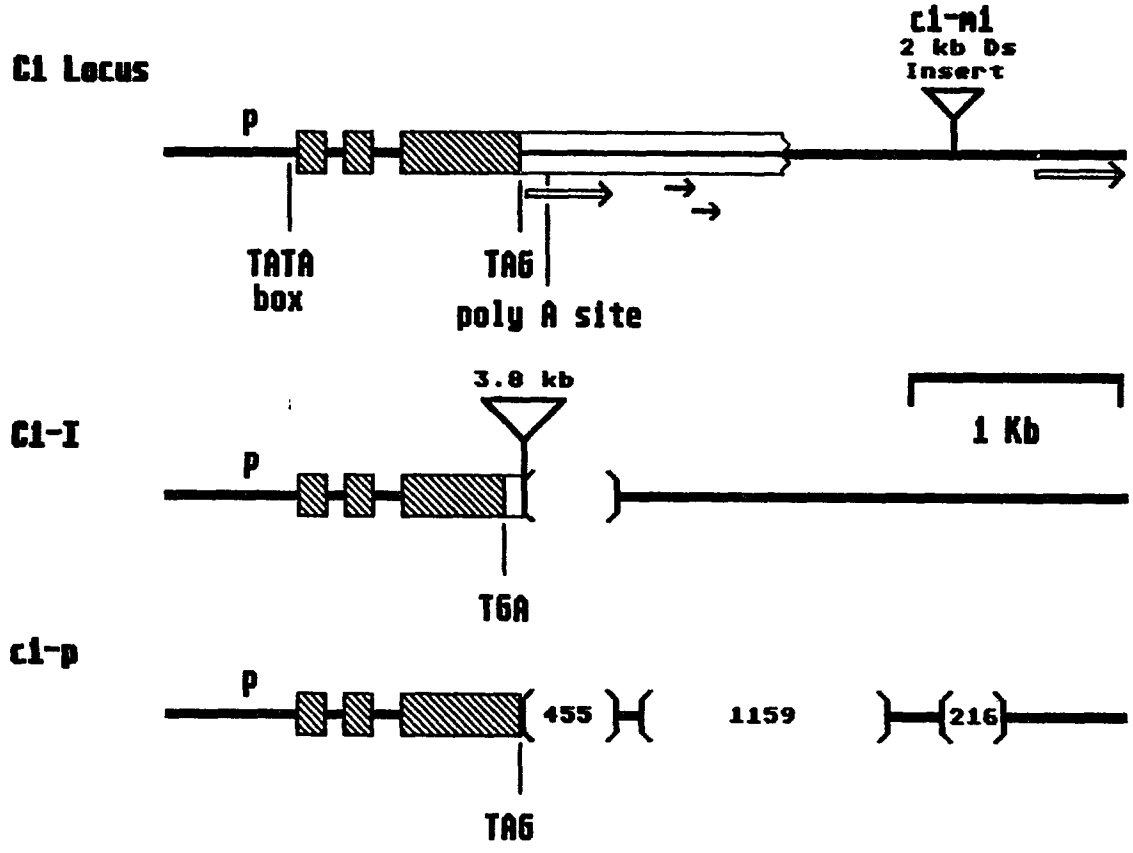


Figure 4.3. Continued

Figure 4.4. A diagrammatic representation of the gene structure of C1, c1-p, c1-m1 and C1-I. The boxes with lines represent the exons. "P" represents the promoter. "TAG" and "TGA" are the translation stop sites. The open boxes represent segments of the transcripts that are not translated. The open arrows show the locations of the large direct duplications. The black arrows represent the smaller direct duplications. The insertions in C1-I and c1-m1 are indicated as triangles. c1-m1 is shown as an insertion in the C1 locus. Deletions in C1-I and c1-p are shown as open brackets

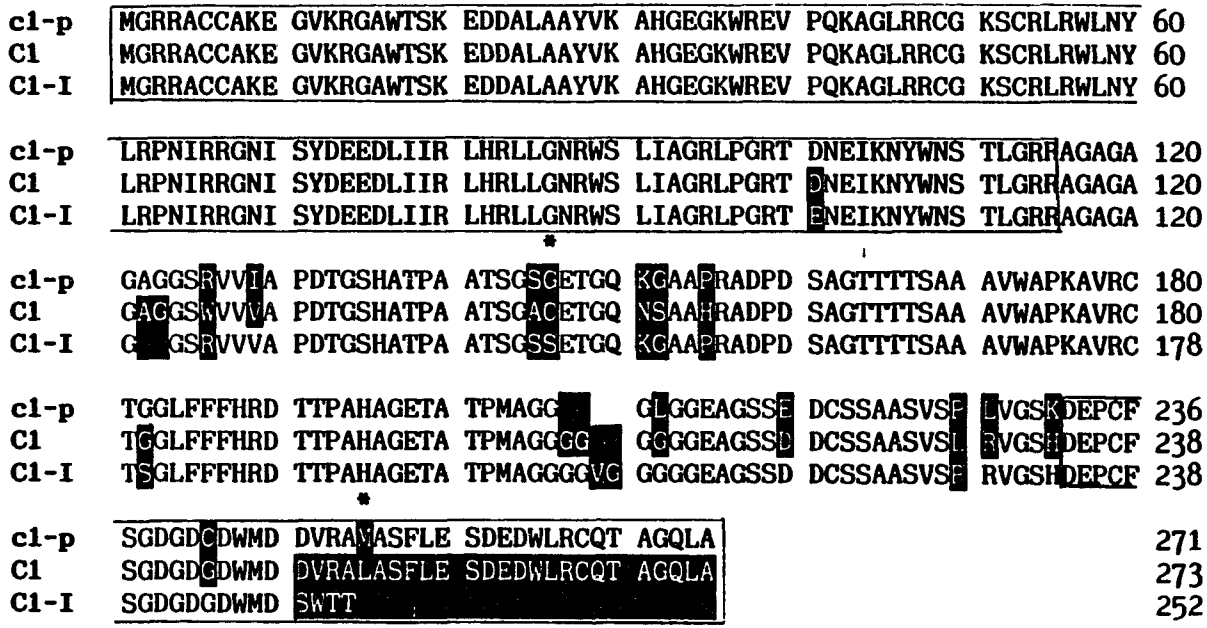

C1 Locus
c1-m1
2 kb Ds
Insert
P
TATA
box
TAG
poly A site
C1-I
3.8 kb
1 Kb
P
TGA
c1-p
P
455
1159
216
TAG

site and one copy of the large duplication are located. The largest deletion (DEL2) is 1159 bp (position 1649 to 2807 in C1) and this removes both copies of a duplicated sequence. The third deletion (DEL3) is 216 bp (position 3057 to 3272 in C1). DEL3 may be significant because the mutant c1-m1 has a Ds transposable element located in this area.

The DNA sequence analysis does not directly show why the c1-p allele is light sensitive. The few alterations in the promoter region seem to be minor and do not create a UV box like the one present in chalcone synthase genes (Schulze-Lefert et al., 1989). The C1 allele not known to have a UV box. It is possible that the deletions in the 3' region may be the reason for light sensitivity because they remove the poly(A) addition site (DEL1) or the site of the Ds insert in the mutant c1-m1 (DEL3). The deletions may remove a 3' regulatory sequence or cause alternative RNA splicing.

The coding regions, as determined by the sequencing of C1 and C1-I cDNA clones (Paz Ares et al., 1987, 1989), were converted into their corresponding amino acid sequences (Figure 4.5). Most of the alterations in c1-p are limited to the area between the acidic and basic domains. There are no amino acid changes in the putative

Figure 4.5. A comparison of the amino acid sequences of C1, c1-p and C1-I. The basic domain (amino-terminus) and the acidic domain (carboxy terminus) are boxed in. Differences between the amino acid sequences, comparing C1 to c1-p or C1-I, are highlighted in black. An asterisk indicates that the sequence of the three alleles is different at that location


c1-p MGRRACCAKE GVKRGAWTSK EDDALAAYVK AHGEGKWREV PQKAGLRRCG KSCRLRWLNY 60
C1 MGRRACCAKE GVKRGAWTSK EDDALAAYVK AHGEGKWREV PQKAGLRRCG KSCRLRWLNY 60
C1-I MGRRACCAKE GVKRGAWTSK EDDALAAYVK AHGEGKWREV PQKAGLRRCG KSCRLRWLNY 60
c1-p LRPNIRRGNI SYDEEDLIIR LHRLLGNRWS LIAGRLPGRT DNEIKNYWNS TLGRRAGAGA 120
C1 LRPNIRRGNI SYDEEDLIIR LHRLLGNRWS LIAGRLPGRT DNEIKNYWNS TLGRRAGAGA 120
C1-I LRPNIRRGNI SYDEEDLIIR LHRLLGNRWS LIAGRLPGRT ENEIKNYWNS TLGRRAGAGA 120
c1-p GAGGSRVVIA PDTGSHATPA ATSGSGETGQ KGAAPRADPD SAGTTTTSAA AVWAPKAVRC 180
C1 GAGGSWVVVA PDTGSHATPA ATSGACETGQ NSAAHRADPD SAGTTTTSAA AVWAPKAVRC 180
C1-I G GSRVVVA PDTGSHATPA ATSGSSETGQ KGAAPRADPD SAGTTTTSAA AVWAPKAVRC 178
c1-p TGGLFFFHRD TTPAHAGETA TPMAGG GLGGEAGSSE DCSSAASVSP LVGSKDEPCF 236
C1 TGGLFFFHRD TTPAHAGETA TPMAGGGG GGGGEAGSSD DCSSAASVSL RVGSHDEPCF 238
C1-I TSGLFFFHRD TTPAHAGETA TPMAGGGGVG GGGGEAGSSD DCSSAASVSP RVGSHDEPCF 238
c1-p SGDGDCDWMD DVRAMASFLE SDEDWLRCQT AGQLA 271
C1 SGDGDGDWMD DVRALASFLE SDEDWLRCQT AGQLA 273
C1-I SGDGDGDWMD SWTT 252

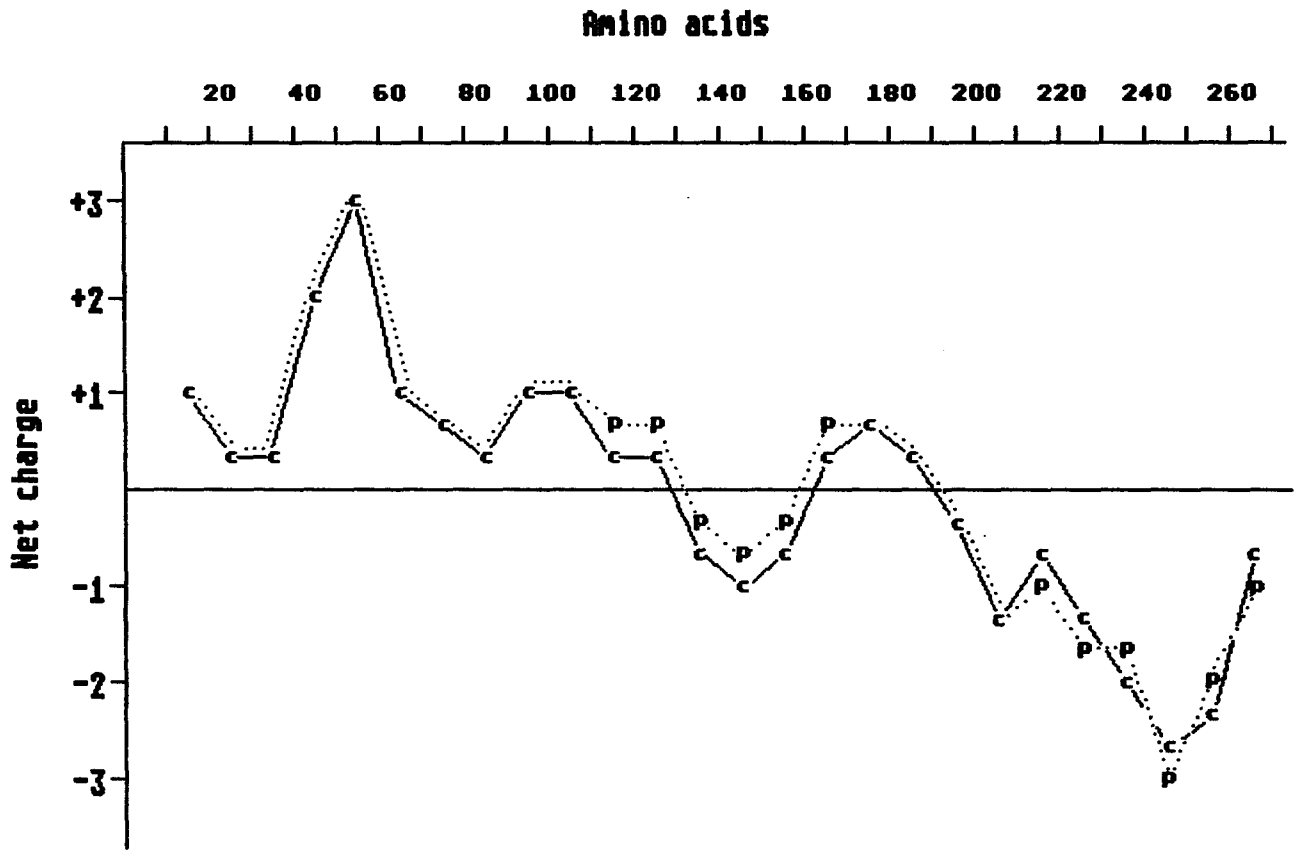
DNA binding domain (basic) and changes in the major acidic domain of c1-p do not decrease its acidity. There is a slight decrease in acidity in the minor acidic domain. This is shown in Figure 4.6 where the charge distribution of the putative proteins encoded by C1 and c1-p are compared. The comparison is done by averaging the net charge over 30 successive amino acids, measured at 10 amino acid intervals. Arginine (R) and lysine (K) are basic and aspartic acid (D) and glutamic acid (E) are considered acidic.

The amino acid alterations in c1-p indicate that the putative protein would be fully functional because the acidic and DNA binding domains are intact. There is no evidence that the light induction nature of c1-p is due to any alterations in the protein structure.

4.1.4. Heteroduplex analysis of c1-p, C1 and C1-I

Because there is no sequence data for the 3' part of the C1-I allele and because c1-p and C1-I share some homology at the DNA sequence level, it was of interest to determine if C1-I had any deletions in the 3' region that might be in common with c1-p. Heteroduplex analysis was conducted on the three alleles (C1 versus C1-I and c1-p; c1-p versus C1-I). The analysis was done at the in house

Figure 4.6. A comparison of the charge distribution of the putative C1 and c1-p proteins. This was done by averaging the net charge over 30 successive amino acids, measured at 10 amino acid intervals (lysine and arginine are basic, and glutamic acid and aspartic acid are considered acidic). The C1 protein is plotted with a solid line and each data point is marked with a "c". The c1-p protein is plotted with a dotted line and the data points are marked with "p" unless they share the same location with the C1 protein and then only "c" is shown


Amino acids
20
40
60
80
100
120
140
160
180
200
220
240
260
Net charge
+3
+2
+1
-1
-2
-3

facilities of the Max-Planck-Institut by Isolde Bertram and Margot Küsters.

Heteroduplex analysis is a rapid but imprecise method to determine if two different DNA fragments have sections of DNA in common. Large deletions or insertions can be detected by this method, but cannot be distinguished from each other.

Phage DNA of the lambda genomic clones for c1-p, C1 and C1-I were used for hybridization. The phages were hybridized in formamide and spread onto nitrocellulose film on copper grids. The grids were platinum shadowed and examined under an electron microscope. Photographs of the heteroduplexes were taken and measurements were made on the single and double stranded DNA.

The original analysis of the C1 and C1-I heteroduplexes, combined with sequence data, showed that C1-I had an insert of approximately 4.2 kb. Heteroduplexes between C1 and c1-p confirmed the sequencing data by revealing the three deletions in c1-p of 450 bp, 1.2 kb, and 200 bp. When the heteroduplexes of c1-p and C1-I were analyzed, a surprising result was found. It was expected from the previous analysis that 4.2 kb, 450 bp, 1.2 kb, and 200 bp deletions/insertions

would be found. Instead, 3.8 kb and 1.2 kb sections of single stranded DNA were detected.

All heteroduplex photographs were reexamined and a mistake in the C1/C1-I heteroduplex was found. Although it was thought that C1-I had a 4.2 kb insert, closer scrutiny revealed that at the same site C1-I had a 3.8 kb insert and a 380 bp deletion. The original observation was an error, because deletions and insertions are both detected as single stranded DNA and both of these phenomenon occurred at the same location in the C1-I allele.

In summary, the heteroduplex analysis indicate that C1-I and c1-p have the same deletion (DEL1) in common (Figure 4.4). In addition, C1-I has a 3.8 kb insertion adjacent to DEL1. The sequence data (Figures 4.2 and 4.3) showed that the first bp of the C1-I insert is the same bp where the c1-p DEL1 starts (position 1087 in C1). This gives further support to the supposition that C1-I and c1-p are closely related to each other. It is possible that when the 3.8 kb insert in C1-I inserted into the proposed progenitor allele of C1-I and c1-p, it simultaneously caused a deletion (DEL1).

DEL3 is not detected in the heteroduplex analysis of C1 versus C1-I or C1-I versus c1-p. This deletion is

rather small, and it is possible that its size and the hybridization conditions prevented it from being detected. Therefore, it is not possible to determine if C1-I has this deletion in common with c1-p.

4.2. Molecular Analysis of c1-m1

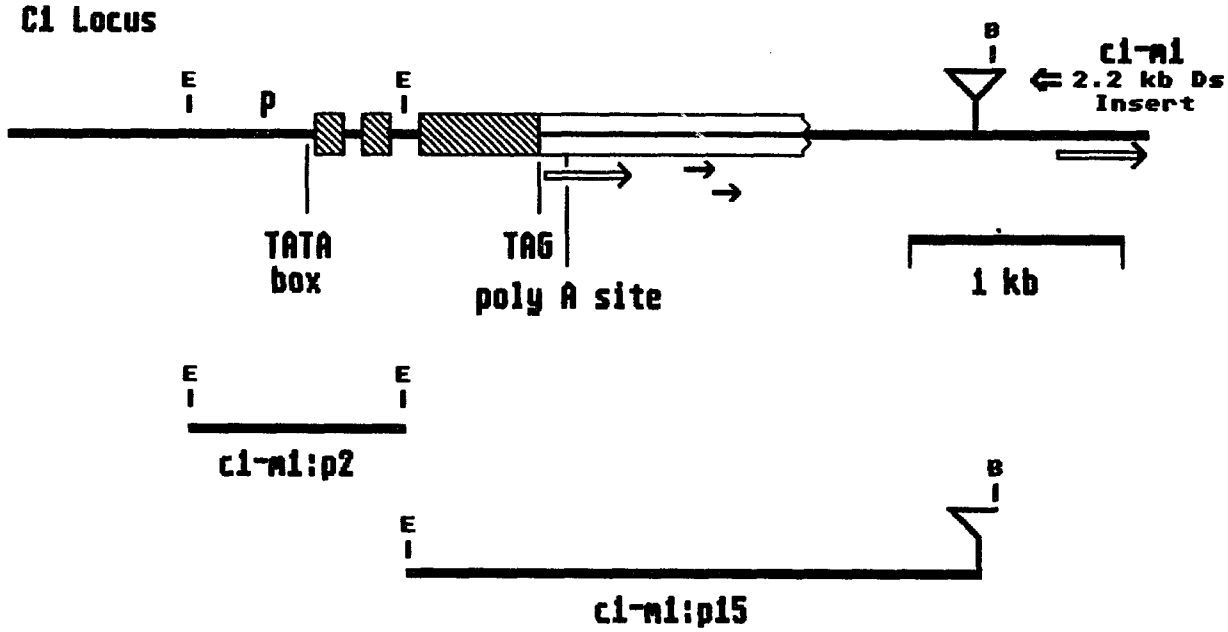
In the preliminary analyses of the c1-m1 allele, it was shown that the transposable element Ds was probably inserted at position 3100 of C1 (Figure 4.4), which is 2025 bp downstream of the translation stop codon (Cone et al. 1986; Förger, 1988). Because c1-p has a deletion at the same site as the Ds insertion in c1-m1, it was important to make a more detailed study of c1-m1 to learn more about the expression and regulation of the wild type allele C1. In addition, c1-m1 is of historical importance because it is the first mutable mutation isolated by McClintock (1948, 1949, 1951). The mutant arose from the transposition of the "standard Ds" into the C1 locus. This Ds element is associated with chromosome breaks and it was of interest to ascertain its structure and to determine if it conformed to a theory postulated by Döring and Starlinger (1984) for Double Ds elements causing chromosome breaks.

4.2.1. Sequencing of the c1-m1 allele

A genomic recombinant phage clone of c1-m1 was obtained from Dr. Karen Cone (University of Missouri). The phages were grown and the DNA extracted as described in Section 3.6. The DNA was double digested with EcoRI and BamHI. A 1.1 kb EcoRI fragment was subcloned into the EcoRI site of the plasmid vector pUC19. This clone (c1-m1:p2) covers a region from the promoter to the middle of the second intron (Figure 4.7). In addition, a 4.2 kb EcoRI/BamHI fragment was cloned into the EcoRI and BamHI sites of pUC19. This clone (c1-m1:p15) contains all the C1 gene from the middle of intron 2 to the site of the Ds insert and most of the Ds element (Figure 4.7).

Deletion derivatives of c1-m1:p2 and c1-m1:p15 were made using Exonuclease III (Sections 3.8.6 and 4.1.2.1), (Henikoff, 1987). Unidirection digestion was achieved by cutting with a restriction enzyme which leaves 3' protruding ends, or by filling 3' recessed ends with α-phosphorothioate deoxynucleoside triphosphates. The clone c1-m1:p2 was digested with PstI and then with BamHI. In contrast, clone c1-m1:p15 was first digested with HindIII and the 3' recessed ends were filled in with α-phosphorothioate deoxynucleoside triphosphates (Henikoff and Hutchinson, 1987). The DNA was then digested with

Figure 4.7. A restriction map of c1-m1. The structure of the C1 locus from Figure 4.4 is shown above the map. The restriction sites are B = BamHI and E = EcoRI. Subclones of the c1-m1 genomic phage clone are listed below the restriction map


C1 Locus
E
P
E
B
ci-m1
⇐ 2.2 kb Ds
Insert
TATA
box
TAG
poly A site
1 kb
E
E
ci-m1:p2
B
E
ci-m1:p15

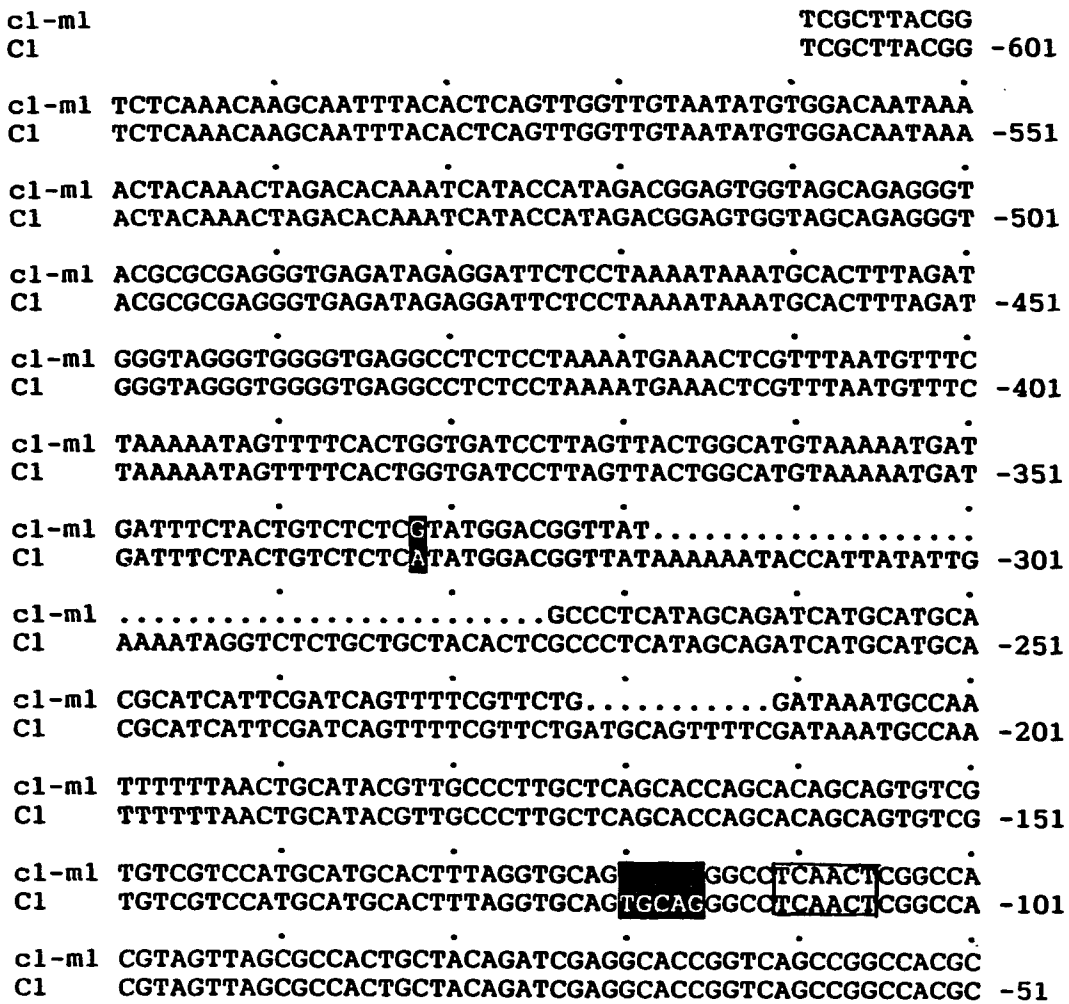
restriction enzyme BamHI. The rest of the protocol and isolation of plasmid DNA were conducted as outlined in Sections 3.4.3 and 3.8.6. The plasmid DNA of the deletion derivatives were used in dideoxy termination sequencing (Sections 3.8.2 and 4.1.2.4). Only one strand was sequenced, because of time restraints.

4.2.2. Sequence of the c1-m1 allele

A comparison of the c1-m1 allele to C1 is shown in Figure 4.8. Although there are several small gaps in the sequence, there are very few differences between the two sequences, indicating that they are closely related. A missing G nucleotide at position 932 in c1-m1 would cause a major frameshift mutation. Because only one strand was sequenced, it is probable that the missing G is due to an artifact called a compression effect. Sequence analysis of the other strand or chemical degradation will be necessary to clarify this result.

The only major difference between C1 and c1-m1 is the presence of the Ds insertion. The Ds element is located at position 3100. The element caused an 8 bp host duplication upon insertion. Because gene function is restored when the Ds element excises, it can be concluded that the minor base pair alterations in the C1 portion of c1-m1 do not have a major effect on gene expression.

Figure 4.8. A comparison of the DNA sequences of C1 and c1-m1. The start and stop sites of translation, TATA box and CAAT box are boxed in. Introns are presented using lower case letters. Differences between C1 versus c1-m1 are highlighted in black. Gaps in the c1-m1 sequence are marked with the symbol ".". The 8 bp that were duplicated by the Ds-cm1 are underlined


c1-m1 TCGCTTACGG
C1 TCGCTTACGG -601
c1-m1 TCTCAAACAAGCAATTTACACTCAGTTGGTTGTAATATGTGGACAATAAA
C1 TCTCAAACAAGCAATTTACACTCAGTTGGTTGTAATATGTGGACAATAAA -551
c1-m1 ACTACAAACTAGACACAAATCATACCATAGACGGAGTGGTAGCAGAGGGT
C1 ACTACAAACTAGACACAAATCATACCATAGACGGAGTGGTAGCAGAGGGT -501
c1-m1 ACGCGCGAGGGTGAGATAGAGGATTCTCCTAAAATAAATGCACTTTAGAT
C1 ACGCGCGAGGGTGAGATAGAGGATTCTCCTAAAATAAATGCACTTTAGAT -451
c1-m1 GGGTAGGGTGGGGTGAGGCCTCTCCTAAAATGAAACTCGTTTAATGTTTC
C1 GGGTAGGGTGGGGTGAGGCCTCTCCTAAAATGAAACTCGTTTAATGTTTC -401
c1-m1 TAAAAATAGTTTTCACTGGTGATCCTTAGTTACTGGCATGTAAAAATGAT
C1 TAAAAATAGTTTTCACTGGTGATCCTTAGTTACTGGCATGTAAAAATGAT -351
c1-m1 GATTTCTACTGTCTCTCGTATGGACGGTTAT.....
C1 GATTTCTACTGTCTCTCATATGGACGGTTATAAAAAATACCATTATATTG -301
c1-m1 .....GCCCTCATAGCAGATCATGCATGCA
C1 AAAATAGGTCTCTGCTGCTACACTCGCCCTCATAGCAGATCATGCATGCA -251
c1-m1 CGCATCATTCGATCAGTTTTCGTTCTG.....GATAAATGCCAA
C1 CGCATCATTCGATCAGTTTTCGTTCTGATGCAGTTTTCGATAAATGCCAA -201
c1-m1 TTTTTTAACTGCATACGTTGCCCTTGCTCAGCACCAGCACAGCAGTGTCG
C1 TTTTTTAACTGCATACGTTGCCCTTGCTCAGCACCAGCACAGCAGTGTCG -151
c1-m1 TGTCGTCCATGCATGCACTTTAGGTGCAG[redacted]GGCCTCAACTCGGCCA
C1 TGTCGTCCATGCATGCACTTTAGGTGCAGTGCAGGGCCTCAACTCGGCCA -101
c1-m1 CGTAGTTAGCGCCACTGCTACAGATCGAGGCACCGGTCAGCCGGCCACGC
C1 CGTAGTTAGCGCCACTGCTACAGATCGAGGCACCGGTCAGCCGGCCACGC -51

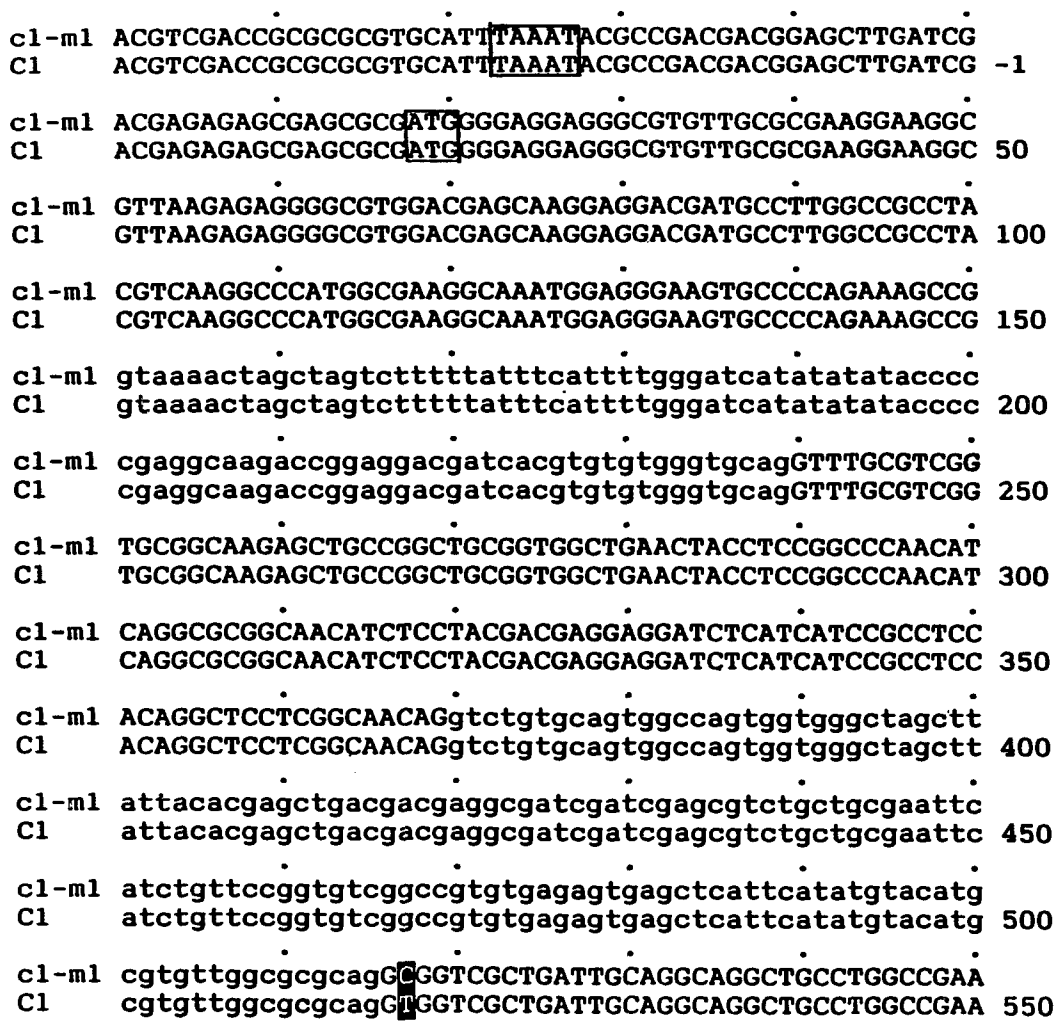


Figure 4.8. Continued

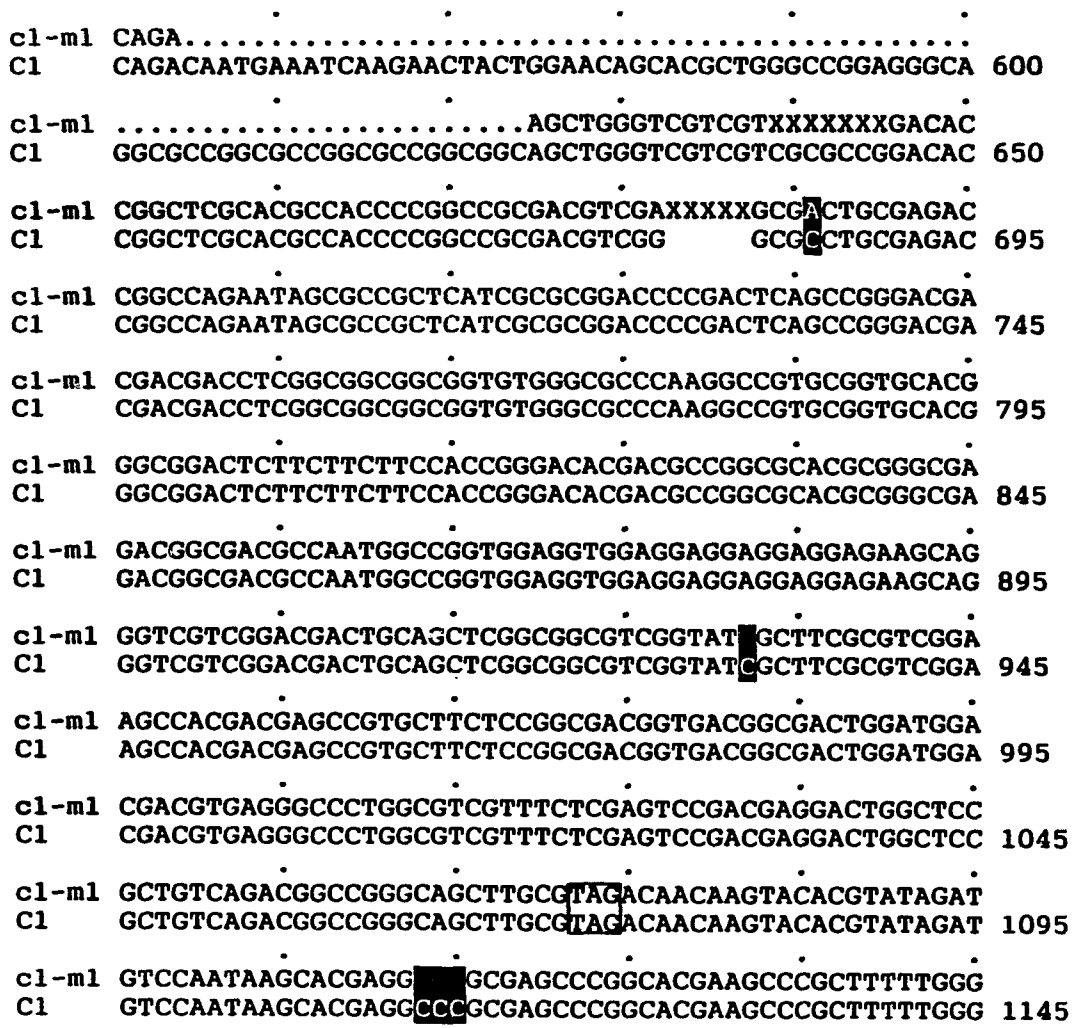


Figure 4.8. Continued

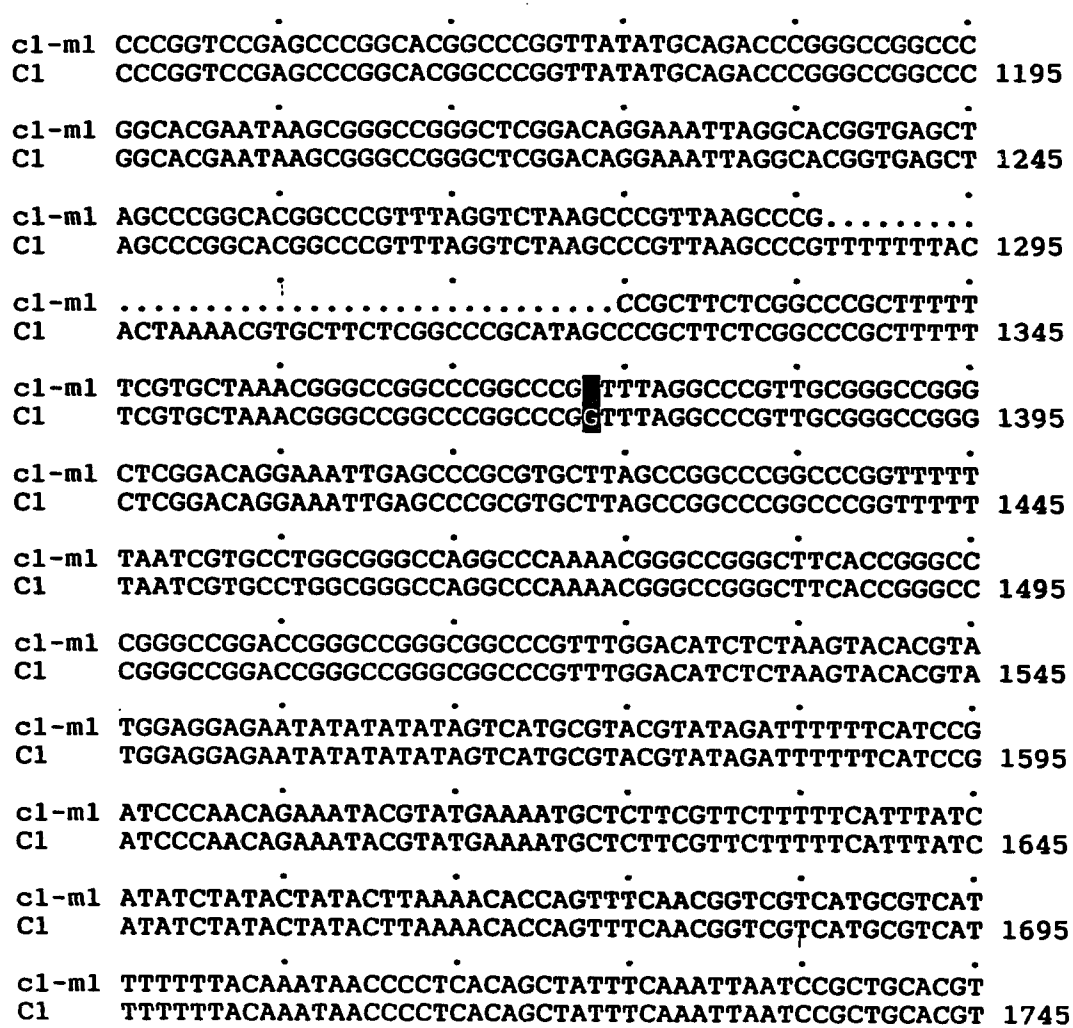


Figure 4.8. Continued

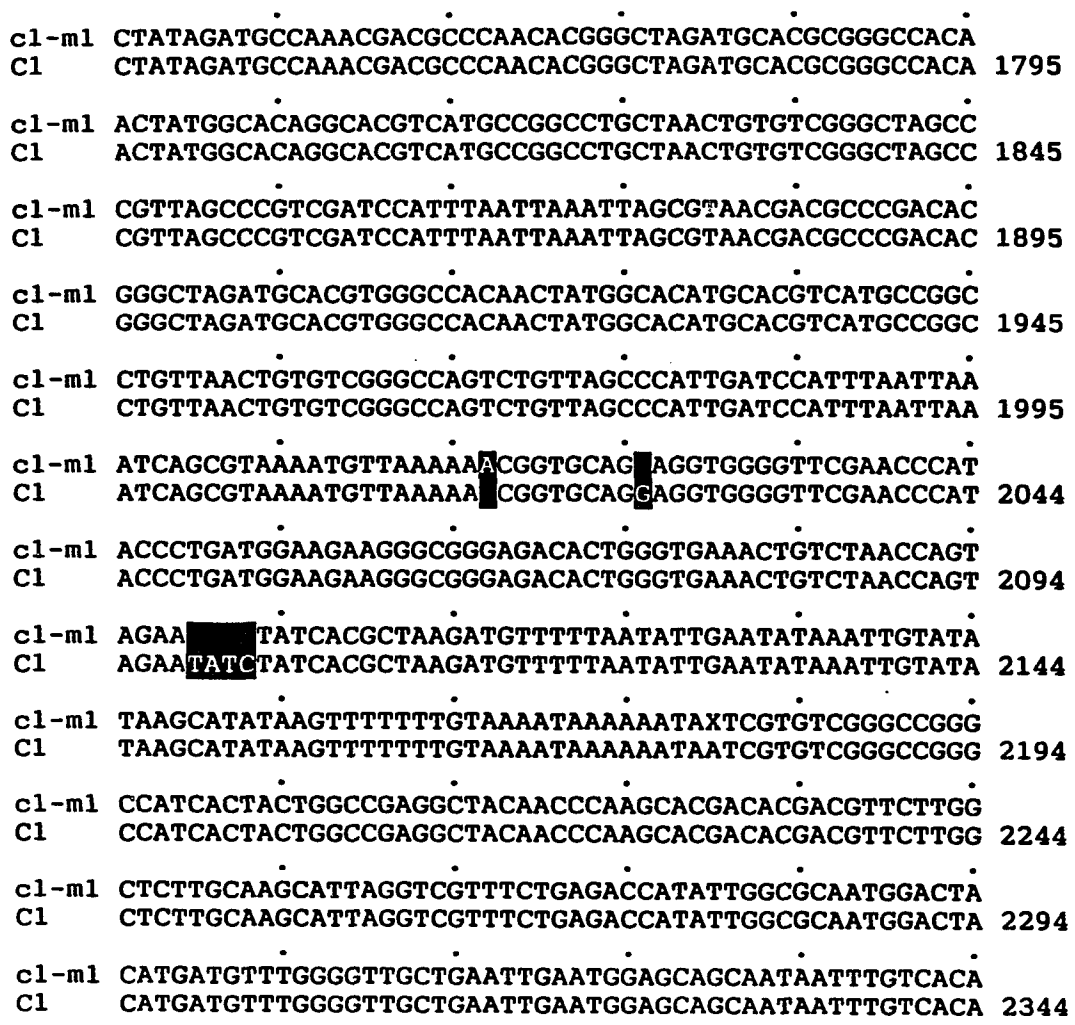


Figure 4.8. Continued

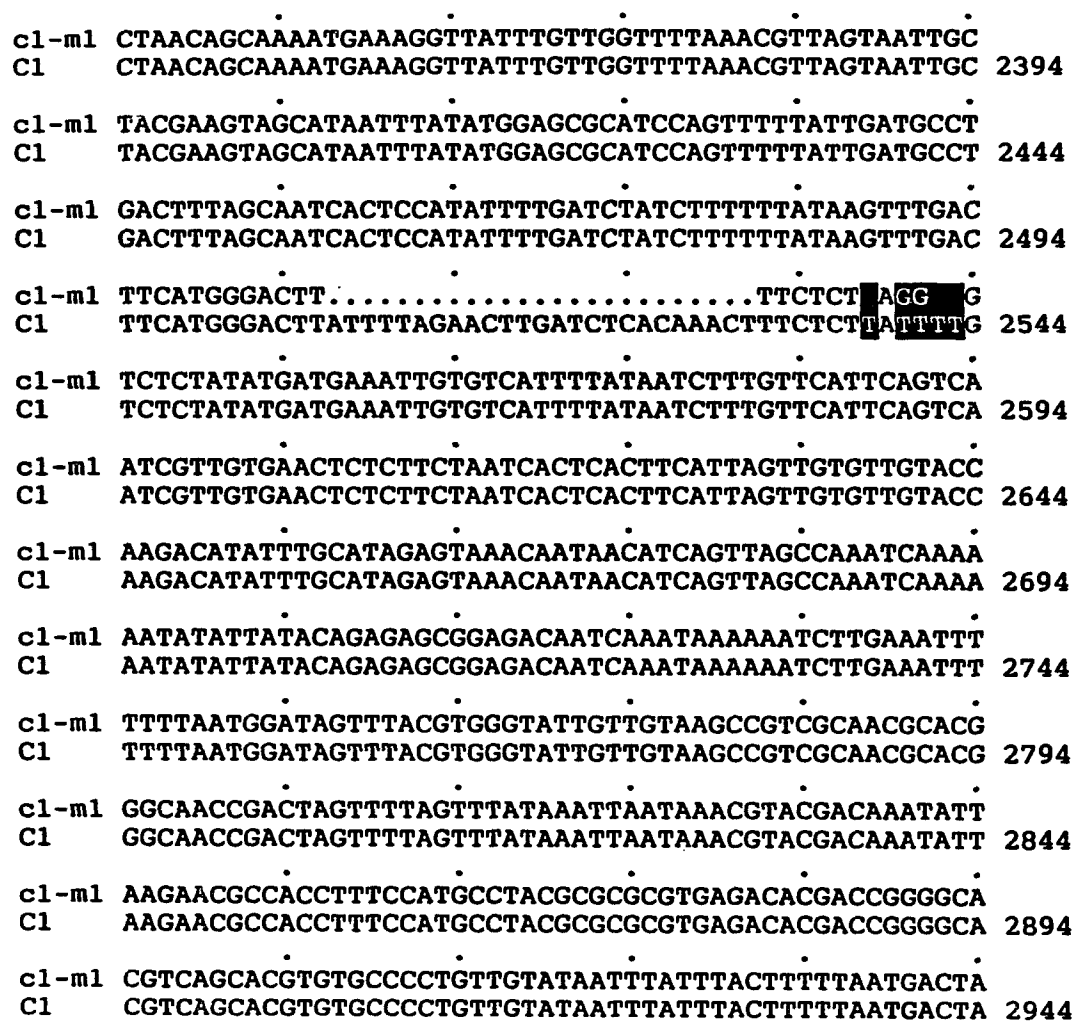


Figure 4.8. Continued

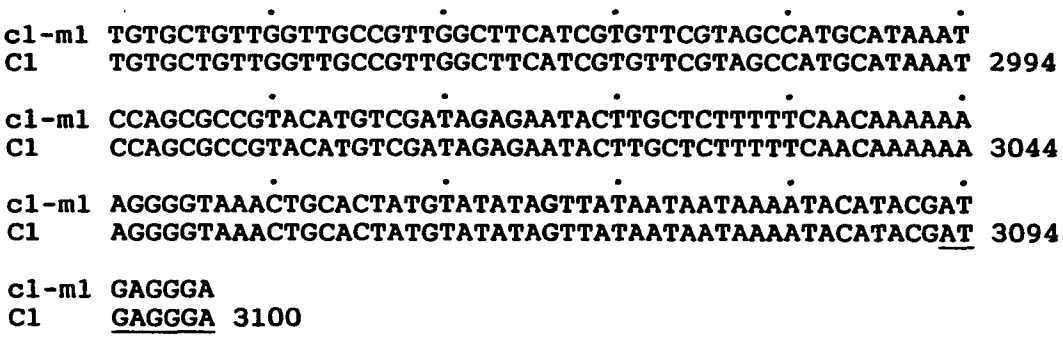


Figure 4.8. Continued

There is no indication why the Ds element would inhibit normal gene expression when it is located 2025 bp downstream of the translation stop site.

4.2.3. Sequence of the Ds element inserted in c1-m1

The sequence of the Ds element inserted in c1-m1 (Ds-cm1) was compared with Ds5933 (Figure 4.9). Ds5933 was isolated from sh-m5933 and is a Double Ds (Döring et al., 1984). A Double Ds element has two exact copies of a deletion derivative Ds element, with one copy inserted in the other. The two copies have opposite orientations. Part of the Ds-cm1 sequence (1937-2067) presented in Figure 4.9 was sequenced by Förger (1988). Although there are gaps in the sequence and one deletion border is poorly defined, the sequence comparison shows that Ds-cm1 is basically homologous to Ds5933, except that Ds-cm1 has only one copy of Ds.

Within Ds5933, there is an 8 bp duplication that was created when the second Ds element inserted into the original Ds element. Ds-cm1 is not a Double Ds element, but it contains this 8 bp duplication. This duplication is identical to that in Ds5933, except that in one copy the terminal G nucleotide was converted to a C. These types of duplications are referred to as transposable element footprints. This footprint indicated that Ds-cm1

Figure 4.9. A comparison of the DNA sequences of Ds5933 and the Ds element inserted in c1-m1 (Ds-cm1). Differences between Ds5933 versus Ds-cm1 are highlighted in black. Gaps in the Ds-cm1 sequence are marked with the symbol ".". The location of the internal Ds element in Ds5933 is marked and highlighted in black. The 8 bp duplication caused by the insertion of the internal Ds element is underlined


Ds-cm1 TAGGGATGAAAACGGTCGGCAACGGTCGGTAAAATACCTCTACCGTTTTC 50
Ds5933 TAGGGATGAAAACGGTCGGCAACGGTCGGTAAAATACCTCTACCGTTTTC 50
Ds-cm1 ATTTTCATATTTAACTTGCGGGACGGAAACGAAAACGGGATATACCGGTA 100
Ds5933 ATTTTCATATTTAACTTGCGGGACGGAAACGAAAACGGGATATACCGGTA 100
Ds-cm1 ACGAAAACGAACGGGATAAATACGGTAATCGAAAACCGATACGATCCGGT 150
Ds5933 ACGAAAACGAACGGGATAAATACGGTAATCGAAAACCGATACGATCCGGT 150
Ds-cm1 CGGGTTAAAGTCGAAATCGGACGGGAXCCGGTATTTTTGTTCGG..... 200
Ds5933 CGGGTTAAAGTCGAAATCGGACGGGAACCGGTATTTTTGTTCGGTAAAAT 200
Ds-cm1 .....TGAAAACATATATTCCAAAACTTAAAAACAAATATAAAAAATTG 250
Ds5933 CACACATGAAAACATATATTCAAAACTTAAAAACAAATATAAAAAATTG 250
Ds-cm1 TAAACACAAGTCTTAATTAAACATAGATAAAATCCATATAAATCTGGAGC 300
Ds5933 TAAACACAAGTCTTAATTAAACATAGATAAAATCCATATAAATCTGGAGC 300
Ds-cm1 ACACATAGTTTAATGTAGCACATAAGTGATAAGTCTTGGGCTCTTGGCTA 350
Ds5933 ACACATAGTTTAATGTAGCACATAAGTGATAAGTCTTGGGCTCTTGGCTA 350
Ds-cm1 ACATAAGAAGCCATATAAGTCTACTAG..... 400
Ds5933 ACATAAGAAGCCATATAAGTCTACTAGCACACATGACACAATATAAAGTT 400
Ds-cm1 .....TGCTCACATCTGGATCACTTAGCATGC 450
Ds5933 TAAAACACATATTCATAATCACTTGCTCACATCTGGATCACTTAGCATGC 450
Ds-cm1 ATAAACTATTACAACCAAGGCTCATCTGTCAACAAACATAAGACACATTG 500
Ds5933 ATAAACTATTACAACCAAGGCTCATCTGTCAACAAACATAAGACACATTG 500
Ds-cm1 CTCATGGAGAGGAGCCACTTGCTACATCTTCATTATTCTTAGAAAATTCT 550
Ds5933 CTCATGGAGAGGAGCCACTTGCTACATCTTCATTATTCTTAGAAAATTCT 550
Ds-cm1 ATTGCGTCTTCATCCTGTTAATACACAAAAATAAGTCAGTTTTGGATAAA 600
Ds5933 ATTGCGTCTTCATCCTGTTAATACACAAAAATAAGTCAGTTTTGGATAAA 600

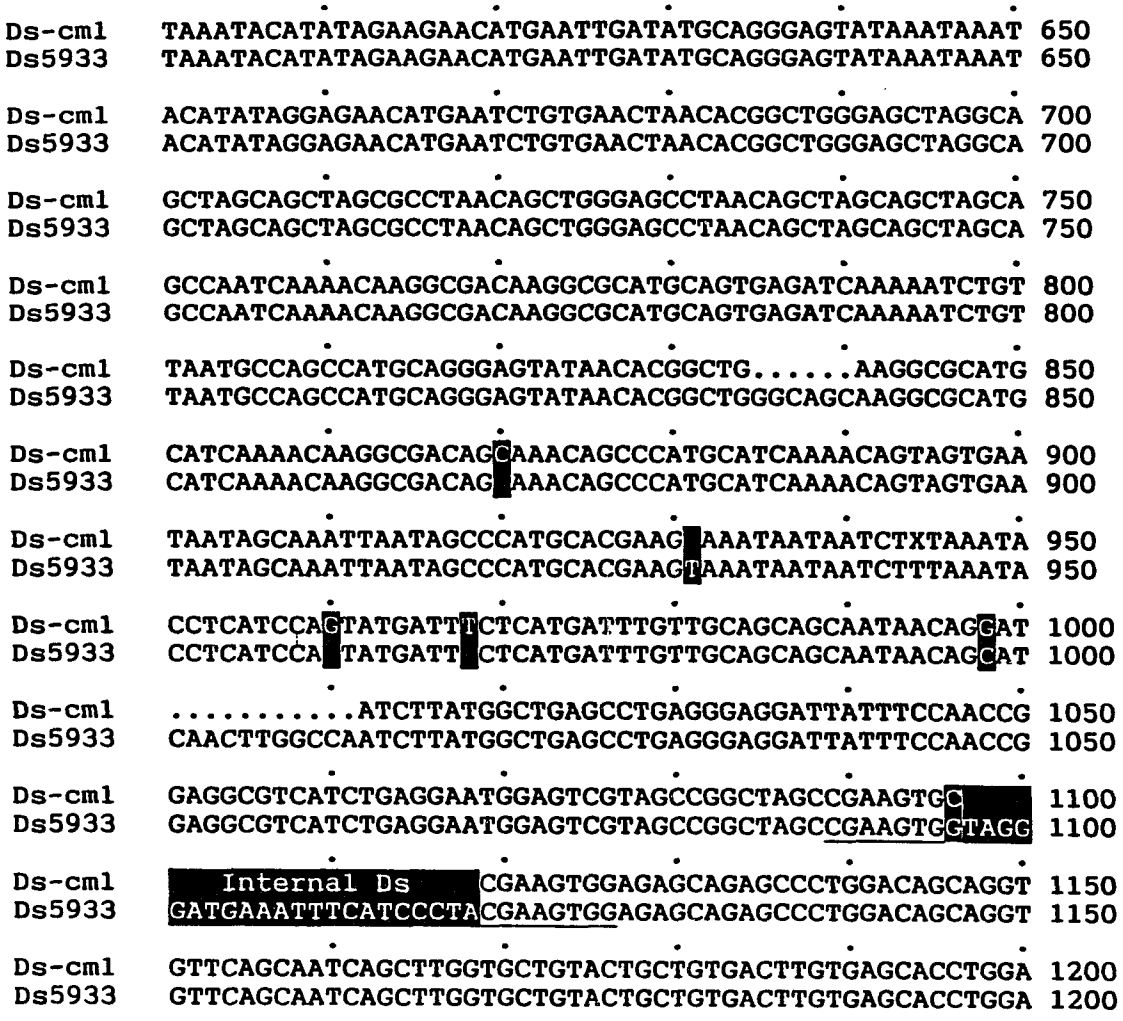


Figure 4.9. Continued

Figure 4.9. Continued

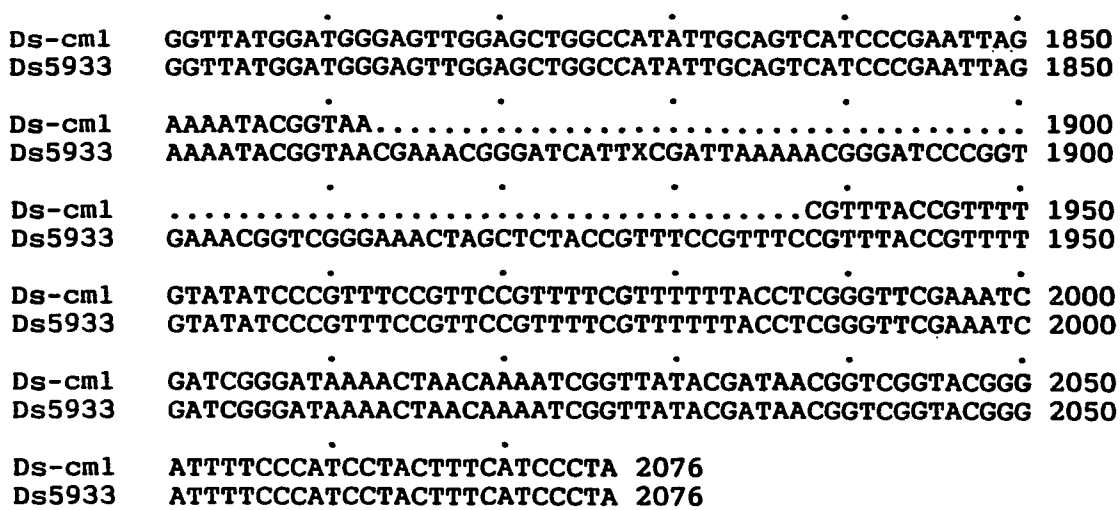


Figure 4.9. Continued

was originally a Double Ds element and at some period in its history the internal Ds element excised out. It is impossible to determine if this event happened after or before Ds-cm1 transposed into the C1 locus.

4.3. Genetic Complementation Test Between c1-p and c1-m1 and Different C1 Mutants

The C1 locus is known to have three transcripts, but it is not clear if they code for the same or different protein product(s). Because of the multiple transcripts and the fact that c1-p and c1-m1 have alterations in the 3' region, it was of interest to determine if other known mutants could complement these two alleles. It is possible that one mutant allele may be deficient for one transcript, while another mutant may be missing a different one. When these two mutants are crossed they might complement each other and the wild type phenotype would be observed in the F1 progeny.

A collection of mutants from Dr. Peter A. Peterson (Iowa State University) were crossed with c1-p and c1-m1 in the field at the Max-Planck-Institut für Züchtungsforschung. Most of the mutants were induced by the insertion of a transposable element, but definitive tests have not been conducted on all the mutants to determine which transposable element is inserted in them.

Due to space and seed limitations, it was not possible to make a diallel analysis of all the mutants. When possible, each mutant was selfed and used as both male and female in crosses with c1-p or c1-m1. The results of this genetic analysis are summarized in Table 4.1.

In all cases, the F1 progeny were spotted (due to excision of the transposable element) or colorless (Table 4.1). These results indicate that none of these mutants could complement c1-p or c1-m1.

A sample of the progeny from all crosses were then germinated in a lighted environmental chamber and evaluated to determine if this stimulus would result in the formation of anthocyanins in the aleurone layer, as it does with homozygous c1-p kernels. Progeny from the crosses to c1-m1 or the self pollinations did not respond to light, except for three mutants. In three instances (c-m11702, c-m32127, and c-m41851), when the mutant plant was crossed onto a c1-m1 female, the progeny did produce some anthocyanin in the aleurone after eight days of the light treatment (Table 4.2). Because color expression occurred later than for homozygous c1-p kernels, it is possible that a different mechanism is responsible for the light response in the three mutants than in c1-p.

Table 4.1. Complementation test of c1-p and c1-m1 with a series of different C1 recessive mutants. Crosses were made using homozygous c1-p and c1-m1 plants as males and females, and the progeny were scored to determine if they were colored

Mutant identification number	Crosses			
	c1-m1		c1-p	
	onto	by	onto	by
c-m6313	-[a]	( )[b]	-	-
c-m11702	-	-	( )	-
c-m804531	( )	( )	-	( )
c-m32127	-	-	-	-
c-m23132	( )	-	-	( )
c-m41963	( )	-	-	( )
c-m55292	( )	-	-	( )
c-m55351	-	( )	-	-
c-m81665	-	( )	-	-
c-m81666	( )	( )	-	( )
c-m55207	-	( )	( )	-
c-m68655	-	-	-	-
c-m55453	-	( )	-	( )
c-m33126	( )	( )	-	-
c-m55607	-	( )	( )	( )
c-m41936	-	-	-	-
c-m63075	( )	-	-	-
c-m63103	( )	( )	-	( )
c-m63103	-	-	( )	-
c-m63120	-	-	-	-
c-m63195	-	-	-	( )
c-m84029	-	( )	-	-
c-m41851	-	-	-	-

[a] "-" indicates there was no complementation.

[b] "( )" indicates no test was made.

Table 4.2. Light induction test of the progeny from the complementation test of c1-p and c1-m1 with a series of different C1 recessive mutants (Table 4.1). Kernels were germinated in a lighted environmental chamber and scored for the production of anthocyanins in the aleurone layer at 2, 2.5, 3, 3.5, and 8 days after the start of germination

Mutant identification number	Crosses			
	c1-m1		c1-p	
	onto	by	onto	by
c-m6313	-[a]	( )[b]	( )	2.0[c]
c-m11702	8.0	-	( )	8.0
c-m804531	( )	( )	2.0	( )
c-m32127	8.0	-	8.0	3.5
c-m23132	( )	-	3.0	( )
c-m41963	( )	-	3.5	( )
c-m55292	( )	-	8.0	( )
c-m55351	-	( )	2.0	2.0
c-m81665	-	( )	( )	3.5
c-m81666	( )	( )	8.0	( )
c-m55207	-	( )	( )	2.0
c-m68655	-	-	3.5	8.0
c-m55453	-	( )	8.0	( )
c-m33126	( )	( )	( )	2.0
c-m55607	-	( )	( )	( )
c-m41936	-	-	2.5	2.0
c-m63075	( )	-	2.0	2.0
c-m63103	( )	( )	2.0	( )
c-m63103	-	-	( )	2.0
c-m63120	-	-	2.0	2.0
c-m63195	-	-	( )	( )
c-m84029	-	( )	2.0	2.5
c-m41851	8.0	-	2.0	2.0

[a] "-" indicates there was no complementation.

[b] "( )" indicates no test was made.

[c] Number of days after the start of germination anthocyanins were detected in the aleurone layer.

Most of the progeny from the crosses with c1-p responded to light, but in several instances the progeny showed little or no response to light (Table 4.2). In most instances there was little difference in the progeny if c1-p was used as a male or a female. It was not possible from this limited analysis to determine if the progeny from the c1-p crosses were light insensitive because of the mutant allele or if the insensitivity was due to environmental factors or the mutant's genetic background.

5. DISCUSSION

5.1. The Importance of the Sequences of C1, C1-I, c1-p and c1-m1

The examination of base pair alterations and their corresponding amino acid changes in different alleles for a particular locus can help define important regions of the locus and facilitate the interpretation of which alterations in a mutant cause the its phenotype. In comparing the sequence of different C1 alleles it is important to remember how the allele was isolated or selected. All the C1 alleles sequenced have one important feature in common and that is that they can produce a protein that is active. Therefore, regions needed for gene expression and functional domains within the protein have been preserved in the alleles.

5.1.1. The structure of c1-p

A comparison of the C1, C1-I and c1-p DNA and amino acid sequences demonstrated that selection for a functional product probably preferentially selected against alterations in the DNA binding domain and the activator domain of the C1 protein. In addition, it seems that most of the promoter sequence is necessary for expression of the C1 locus. The sequences of several old null c1 alleles are needed to demonstrate that prevention

of alterations in these areas is a functional constraint, and not that some other biological mechanism is preferentially preserving them. In contrast, the numerous alterations in the 3' flanking region of these mutants indicates that the same level of preservation in this area is not needed to produce a functional product.

The sequence comparison in Figure 4.2 does not explain the phenotype of c1-p, but by the process of elimination several structural differences can probably be ruled out as being the main factor(s). The sequence alterations in the coding region do not seem to change important regions of the protein Therefore, they probably are not the cause of the c1-p phenotype. In the promoter region, the 5 bp deletion in c1-p at position -123 of the C1 allele (Figure 4.2) is also present in the c1-m1 allele (Figure 4.8). Presence of the deletion in both c1-p and c1-m1 indicated that this deletion is not responsible for the light inducible nature of c1-p. In contrast, the 5 bp deletion at position -152 of the C1 allele (Figure 4.2) is unique to the c1-p allele. Because the promoters of all the C1 alleles sequenced are highly conserved, this small deletion may be a crucial component of c1-p. Sequence comparison of this region to the UV boxes detected by Schulze-Lefert et al. (1989) indicated that this deletion

did not create or destroy a UV box. This observation is not surprising because c1-p is stimulated by red light (Hsu, 1970). It is possible that this deletion did create a sequence that a light induced transcriptional activator could bind too. In vivo DNA footprint analysis could be used to detect if a protein binds to this region and if this protein is light induced.

The most noticeable alterations in c1-p, and therefore the most suspect, are the three major deletions in the 3' region (Figures 4.3 and 4.4). DEL1 is present in c1-p and C1-I, therefore, it is tempting to consider this deletion as unimportant because C1-I has normal expression. On the other hand, this mutation could be very important because it removed 455 bp just after the translation stop site, including the poly(A) site of the 1.1 kb cDNA clone. The elimination of this area is not important in C1-I because its transcription terminates within the 3.8 kb insertion, which is located before DEL1. DEL2 and DEL3 remove sequences that are probably necessary for the 1.6 and 2.5 kb transcripts. THe deletions could be important if the 1.6 and 2.5 kb transcripts are essential for normal C1 expression. DEL3 also eliminates the sequence where Ds-cm1 is located, which may be a regulatory region (Section 5.1.3). Because these deletions remove sections

of DNA that seem to be important, their significance cannot be disregarded. However, the predicted C1 protein coding region is intact, and the presence of the deletions in the 3' part of the c1-p allele suggest that only one protein product from the C1 locus is needed for anthocyanin biosynthesis.

5.1.2. Molecular model for the c1-p phenotype

It is unlikely that a single mutation in c1-p eliminated normal gene expression and caused it to be light inducible. A more probable explanation is that all C1 alleles are light inducible, but phenotypically this is impossible to detect in wild type C1 kernels because anthocyanins are already present. In C1-I kernels, anthocyanin biosynthesis would still be suppressed by the C1-I protein. Because other loci involved in anthocyanin biosynthesis are light inducible (Dangl et al., 1989), this hypothesis is not unreasonable. Therefore, the most likely scenario is that an alteration in c1-p prevents normal gene expression and this defect can be overcome to some extent by exposing the kernels to light.

The simplest hypothesis for the c1-p phenotype is that the 5 bp deletion in the promoter at position -123 of C1 (Figure 4.2) drastically decreases normal gene expression. In the presence of light, an additional regulatory

mechanism is used to enhance gene expression and this increase is great enough to stimulate anthocyanin biosynthesis in homozygous c1-p kernels.

An alternative hypothesis is that deletions in the 3' part of c1-p decrease the production of mRNA, while light would increase mRNA level, as proposed in the previous hypothesis. There are several reasons why these deletions would decrease mRNA production. One reason is that DEL1 removes a known poly(A) site.

For either hypothesis, the transcription mechanism must use another poly(A) site. If this new poly(A) site is weak, as in the case of human α2 globin gene (Higgs et al., 1983; Whitelaw and Proudfoot, 1986) or the DNA tumor virus polyoma (Acheson, 1984; Lanoix and Acheson, 1988), C1 specific mRNA production could be drastically lower than in the wild type gene. A mutation in the normal poly(A) site in the α2 globin gene decreases mRNA production by 80% (Higgs et al., 1983; Whitelaw and Proudfoot, 1986). By increasing the level of transcription, by light induction, the total amount of C1 specific mRNA could be increased to a level where anthocyanin biosynthesis can be stimulated.

5.1.3. The structure of c1-m1

The sequences of C1 and c1-m1 were found to be very similar, which indicated that no other alteration besides the Ds-cm1 insert probably contributes to the mutant phenotype of c1-m1. This result is not surprising because the insertion of Ds-cm1 into a functional wild type C1 allele is a recent event and excision of the element restores normal gene expression (McClintock, 1948, 1949, 1951).

The observed location of Ds-cm1 demonstrated that at least a portion of the 3' part of the gene is important. The reason for its importance is not obvious. It is feasible this area is a regulatory region of the C1 locus. Regulatory regions in the 3' end of a gene have been reported for other loci, such as the Adh2 gene of Drosophila mulleri (Fischer and Maniatis, 1986) and the β-globin and histone (H5) genes from chicken (Choi and Engel, 1986; Trainor et al., 1987). In vivo DNA footprint analysis may reveal that a protein normally binds in this area and Ds-cm1 prevents it from binding.

An alternative hypothesis is that Ds-cm1 could affect transcription at the mRNA level. Northern analysis of c1-m1 homozygous kernels without Ac has not been conducted to determine what, if any, transcripts are produced by

c1-m1. Several models could be proposed to explain how Ds-cm1 may interrupt transcription. One model is that Ds-cm1 blocks normal processing of the primary transcript. In this model, transcription reads into Ds-cm1 and the presence of the element within the nascent RNA inhibits splicing. A second model is that Ds-cm1 could inhibit transcription by causing premature termination of transcription. In this model the RNA would not be polyadenylated and thus unstable (Proudfoot and Whitelaw, 1988). As a result, very little, if any, C1 specific mRNA would be produced in c1-m1 kernels.

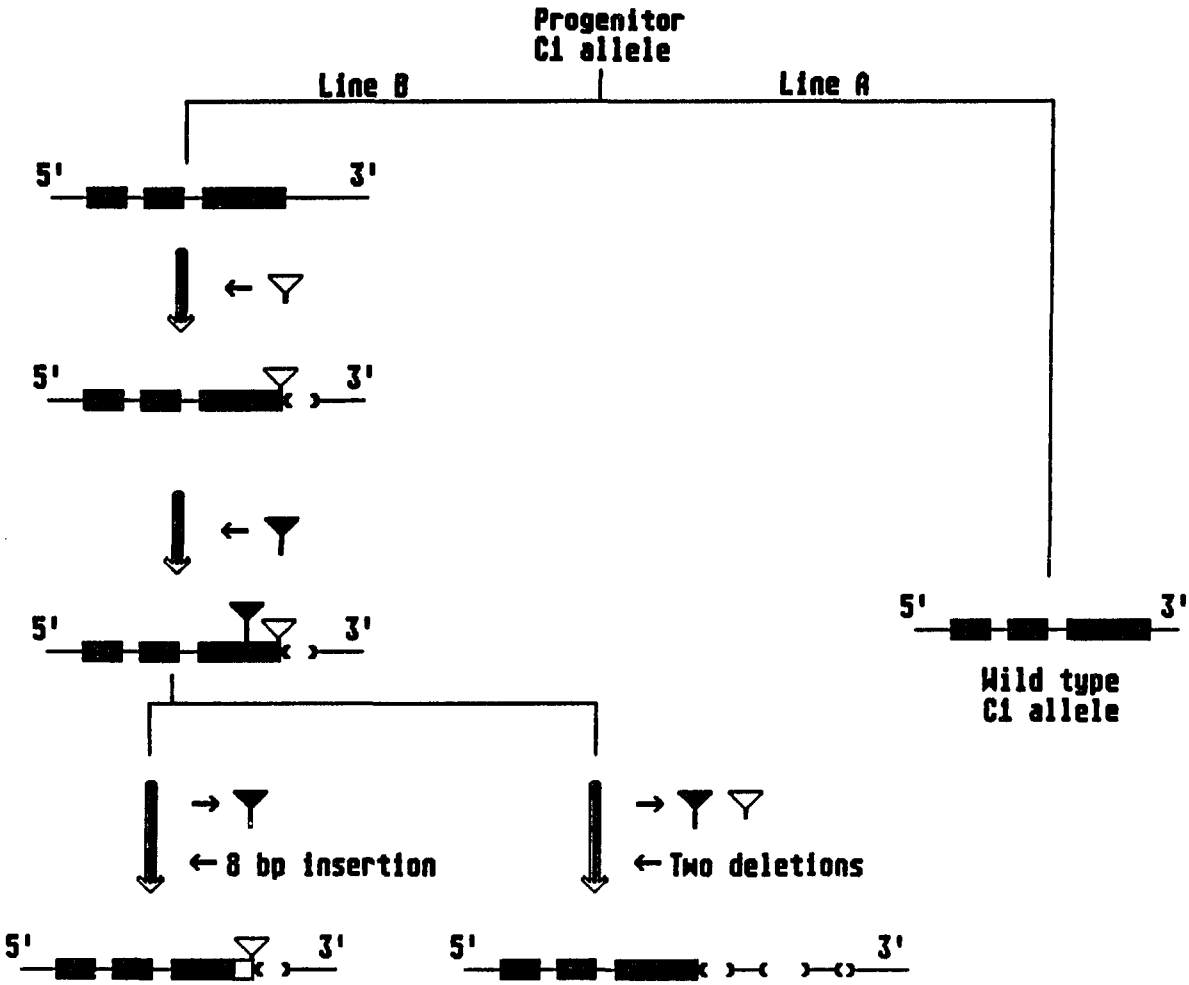
It is not possible from my study to determine the reason why c1-m1 and c1-p are mutant alleles. There are so many alterations in c1-p that it is difficult to ascertain which ones are responsible for its mutant phenotype. The Ds insertion in c1-m1 indicates that the 3' region where it is inserted is important, but does not explain why this insertion would alter gene expression. Molecular evaluation by Northern analysis and cDNA cloning may provide further insight into the nature of these mutant alleles and the regulation of gene expression of the C1 locus.

5.2. The Evolution of the Different C1 Alleles

By comparing the sequence of the known C1 alleles, it is possible to determine a chronological order for the different alterations and to assess which alleles are more closely related.

The number of major alterations that were found to be in common between C1-I and c1-p demonstrated that these two alleles are more closely related to each other than they are to C1. A logical explanation of how these alleles originated would be to propose that there was probably a common progenitor to all C1 alleles. From this progenitor, at least two major divergent lines may have evolved. One line (line A) gave rise to C1 (Paz-Ares et al., 1987) and the other (line B) to C1-I and c1-p (Figure 5.1). In line B, the insert found in the C1-I allele inserted into its progenitor locus and simultaneously caused a deletion (DEL1) (Figure 5.1). At a later date, another transposable element probably inserted into the locus (position 871 in Figure 4.2). On one occasion this new insert excised and deleted 6 bp. This deletion occurred in the c1-p divergent line. In another instance, the transposable element excised out and left behind an additional 6 bp. C1-I was derived from this line. Base pair differences between c1-p and C1-I would have occurred

Figure 5.1. The hypothetical evolution of C1, C1-I and c1-p. The three exons of the C1 locus are shown as black boxes. The large insert found in the C1-I allele (depicted as an open triangle) inserted into and simultaneously caused a deletion in the progenitor allele. Later, a transposable element (depicted as a black triangle) inserted into the locus. When this element excised, it left behind a 6 bp duplication in the C1-I progenitor allele, or it removed 6 bp in the allele that gave rise to c1-p. During the time period when this transposable element excised from the c1-p progenitor allele, the first insertion was also removed and two deletions occurred in the 3' end of the allele. Probably the last major alteration to occur was the 8 bp insertion into the C1-I allele, which caused a frameshift mutation (shown as an open box)


Progenitor
Ci allele
Line B
Line A
5'
3'
5'
3'
5'
3'
5'
3'
Wild type
Ci allele
← 8 bp insertion
← Two deletions
5'
3'
5'
3'

during this point of divergence. In addition, two additional deletions (DEL2, DEL3) in the c1-p line would have taken place and the insert at position 871 would then have transposed out of the progenitor, which gave rise to c1-p. This insert did not excise in the progenitor line that gave rise to C1-I. Naturally, this is only a hypothesis for the evolution the different C1 alleles.

The 8 bp insertion in C1-I that caused a frameshift mutation and turned the allele into a transcriptional suppressor (Section 2.3.4, Figure 5.1) (Paz-Ares et al., 1989) probably occurred very recently because there are no base pair alterations in the sequence after the new translation stop site. Because this region constitutes the putative activator domain of the C1 locus, its sequence would be expected to be relatively conserved over time. After the frameshift mutation, there would be no natural selection against mutations occurring after the stop site. The lack of base pair alterations in this area indicated little time has lapsed to allow mutations to occur in this region.

The sequence analysis of C1 and c1-m1 indicated that the two alleles are practically identical, which indicated that they are closely related. Besides the Ds insert, the only important alteration in c1-m1 is the 5 bp deletion in

the promoter (position -123 in Figure 4.8). This deletion is also present in c1-p, and C1-I has a sequence alteration at the same location. There is no simple explanation as to why these alterations have occurred at this site, but several DNA sequence altering events were necessary to create these differences.

5.3. Chromosome Breakage and the Ds Element Inserted in c1-m1

The sequence of the Ds element inserted in the c1-m1 allele (Ds-cm1) demonstrated that this element is related to Ds5933, except that it does not contain a second copy of the Ds element inserted within itself (Döring et al., 1984; Weck et al., 1984). In Ds5933, the insertion of the internal Ds caused an 8 bp footprint. This footprint, with one nucleotide altered, is also present in Ds-cm1. This nucleotide alteration was probably caused when the internal Ds element was excised (Saedler and Nevers, 1985). The presence of the same footprint indicated that Ds-cm1 was originally a Double Ds element and that sometime in its history the internal Ds element was excised. It is impossible to determine if this excision event took place before or after Ds-cm1 inserted into the C1 locus.

The original isolate of c1-m1 (McClintock, 1948, 1949, 1951) exhibited a high rate of chromosome breakage and had a low somatic reversion rate (c1-m1 -> C1). Some germinal derivatives of c1-m1 had the reverse phenotype, a low rate of chromosome breakage, and a high reversion rate (McClintock, 1948, 1949, 1951). It is not known if the germplasm used to isolate c1-m1 represented the original mutant state or a derivative. Determination of reversion rate is a relative evaluation and at the time of my study, there was no available picture of the original c1-m1 state to use as a control. The seed source used for the cloning of c1-m1 seems to have a high reversion rate, indicating that the c1-m1 allele cloned and sequenced is a derivative and not the original isolate. The rate of chromosome breakage was not determined because no control was available.

Döring and Starlinger (1984) have proposed that Double Ds elements are the factors responsible for the chromosome breakage described by McClintock (1947) (Section 2.3.2). My examination of c1-m1 would partially contradict this hypothesis because Ds-cm1 is not a Double Ds element. Instead the following modification of their hypothesis would explain the molecular and genetic data of c1-m1. Because the original c1-m1 allele had a high rate of

chromosome breakage it, is possible that initially Ds-cm1 was a Double Ds element. In this instance, when the transposase from Ac interacts with a Double Ds element, chromosome breakage, as diagramed in Figure 2.4, occurs more frequently than an excision event that would restore normal gene function (Figure 5.2). In contrast, the derivative alleles of c1-m1, which have a low rate of chromosome breakage and high reversion rate, would represent examples where the internal Ds of the Double Ds element excised, resulting in a structure similar to Ds-cm1. In these cases, it is proposed the internal Ds element transposed to a site nearby on the same chromosome arm. Because Ds-cm1 represents a normal deletion derivative of Ac, it is logical to assume that when the Ac transposase interacts with it normal excision takes place (c1-m1 -> C1) (Figure 2.3). An alternative is that the transposase could simultaneously interact with Ds-cm1 and a nearby Ds element, in a similar fashion as proposed for Double Ds elements (Figure 2.4), which would result in chromosome breakage (Figure 5.3). Because these two Ds elements are separated from each other, the occurrence of this event would be expected to be much lower than the normal excision rate.

Figure 5.2. A diagram of the excision of a Double Ds element that does not cause chromosome breaks and restores gene activity. The black boxes indicate the 8 bp host direct duplication. The open arrows indicate the element's 11 bp inverted repeats. Upon excision the inverted repeats are aligned and the DNA is cleaved at the site indicated with the black arrows. The Ds element is removed and the chromosomal fragments (1-2 and 3-4) are ligated together

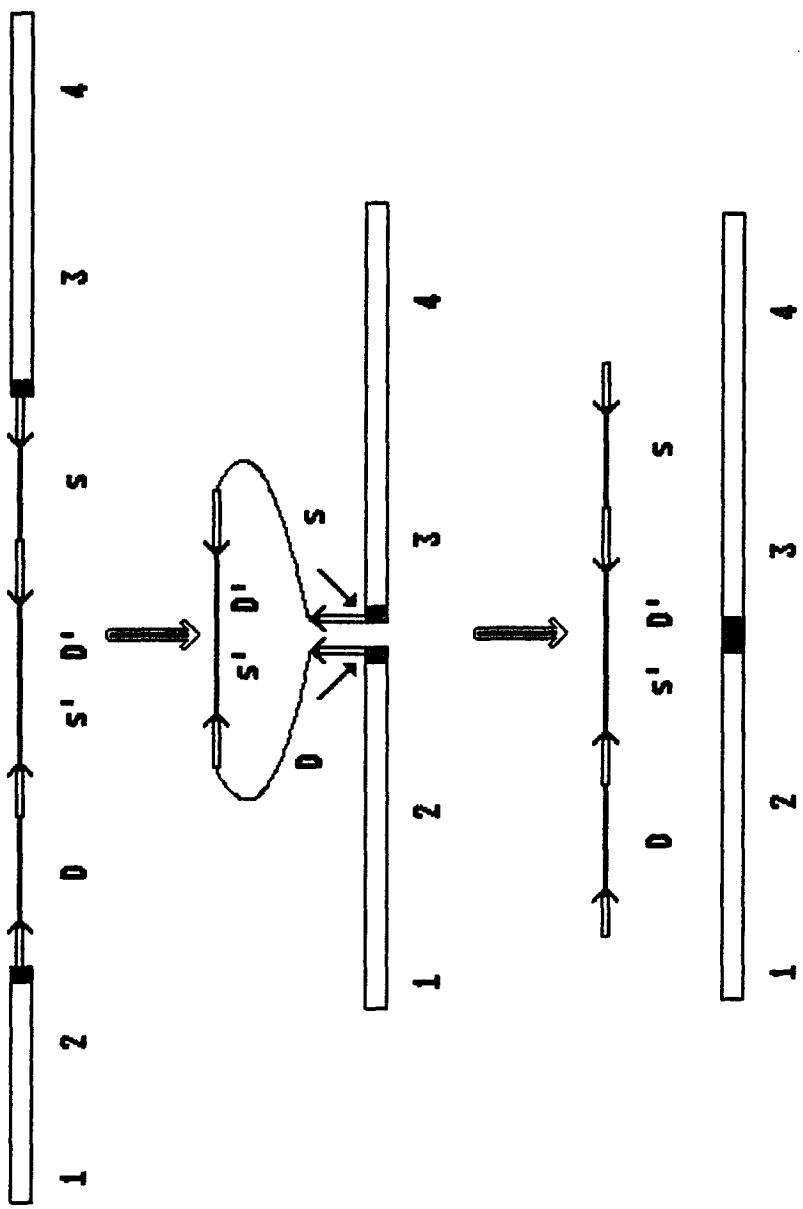
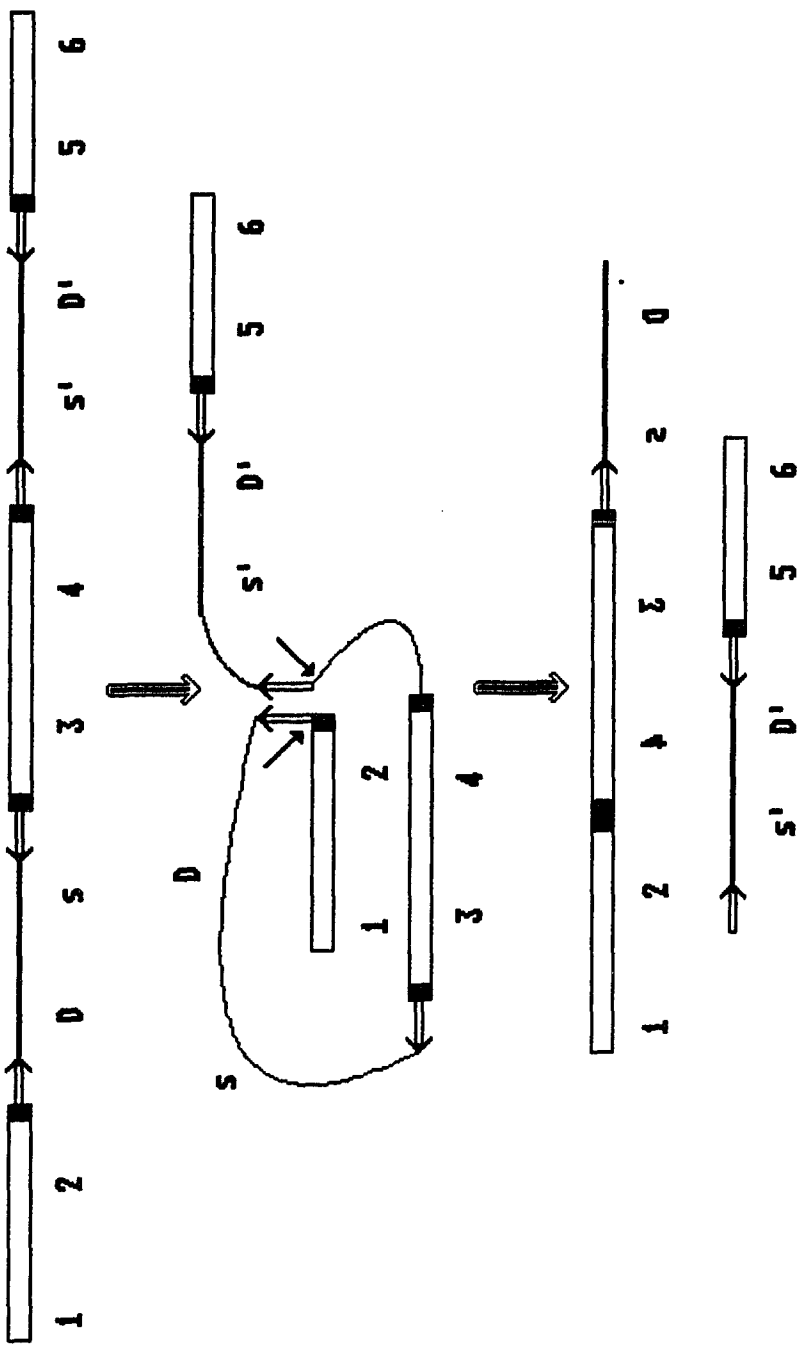

1
2
D
S' D'
S
3
4
D
S' D'
S
1
2
3
4
D
S' D'
S
1
2
3
4

Figure 5.3. A diagram of chromosome breakage caused by two separated Ds elements. The general mechanism is the same as in the case of a Double Ds element (Figure 2.4), except that the Ds elements are separated. One Ds element is inserted in the gene of interest and a second Ds element, indicated as "s' D'", is inserted in an opposite orientation at a nearby site. The black boxes indicate the 8 bp host direct duplication. The open arrows indicate the element's 11 bp inverted repeats. Upon excision, one inverted repeat from each Ds element is aligned, and the DNA is cleaved at the site indicated with the black arrows. Chromosomal fragment 3-4, with the Ds element still attached, is ligated in the opposite orientation to fragment 1-2. The chromosomal segments are not rejoined thus resulting in a chromosome break


1
2
D
S
3
4
S'
D'
5
6
S
1
2
D
3
4
S'
D'
5
6
1
2
Σ
4
2
D
S'
D'
5
6

Although *c1-m1* originally exhibited chromosome breakage at the *C1* locus, probably due to a *Double Ds* element (*Ds-cm1*), this may not be true for the derivative alleles. It is conceivable that the low rate of chromosome breakage observed in the *c1-m1* derivative alleles is caused by a second *Double Ds* element nearby the *C1* locus.

To determine if these hypotheses are correct, genomic Southern analysis is needed to test if the original *c1-m1* isolate contains a *Double Ds* element. More elaborate studies are then needed to determine if *Ds-cm1* can cause chromosome breaks.

6. APPENDIX

The sequence of the C1 locus from Paz-Ares et al. (1987). The CAAT box, TATA box, translation start site, stop codon and poly(A) addition sequence are highlighted in black. Introns 1 (88 bp) and 2 (144 bp) are given in small case letters.

TATCAACCTC -1051

CTGTGTTATT TTTAGTGACG GTTTCTTAAA AAACACCACT AGAAATCGTA -1001

TTTTTATAGG TGGTTCCTTA AGAAAACTGC ATGCAGAAAT CCATGACGGT -951

TTTCTTAAGG AACCGTATGT AGAAATACGA TTTCTAGTGA CGATCTTCTT -900

AAGGAAACCA CCACTAAAAA TTATTTTTAT CCTTAATTTT CGAGTTTTTC -851

AAACGATCTC GTATGATGAA ACCATCAAAA TAAAAGTTGT ACATCTCTAA -801

AAGTTATGAA AATTTGTAGT TAACAACTTT TTTATTTGAA CTCATTTTGG -751

TTCTCAAAAA TTGCATCTAA ATTTGTCAAA TTTAAAATTC AAATTTTCCA -701

AACGACCTCG GATGAAAAAA GTGTCAAAAT GAAAGTTGTA GAACTTCAAA -651

AGTTATTCAA CTTTGTAGTC GACTATCTTT TTATTTGAAT TCGCTTACGG -601

TCTCAAACAA GCAATTTACA CTCAGTTGGT TGTAATATGT GGACAATAAA -551

ACTACAAACT AGACACAAAT CATACCATAG ACGGAGTGGT AGCAGAGGGT -501

ACGCGCGAGG GTGAGATAGA GGATTCTCCT AAAATAAATG CACTTTAGAT -451

GGGTAGGGTG GGGTGAGGCC TCTCCTAAAA TGAAACTCGT TTAATGTTTC -401

TAAAAATAGT TTTCACTGGT GATCCTTAGT TACTGGCATG TAAAAATGAT -351

GATTTCTACT GTCTCTCATA TGGACGGTTA TAAAAAATAC CATTATATTG -301

AAAATAGGTC TCTGCTGCTA CACTCGCCCT CATAGCAGAT CATGCATGCA -251

CGCATCATTC GATCAGTTTT CGTTCTGATG CAGTTTTCGA TAAATGCCAA -201

TTTTTTAACT GCATACGTTG CCCTTGCTCA GCACCAGCAC AGCAGTGTCG -151

TGTCGTCCAT GCATGCACTT TAGGTGCAGT GCAGGGCCTC AACTCGGCCA -101

CGTAGTTAGC GCCACTGCTA CAGATCGAGG CACCGGTCAG CCGGCCACGC -51

ACGTCGACCG CGCGCGTGCA TTTAAATACG CCGACGACGG AGCTTGATCG -1

ACGAGAGAGC GAGCGCGATG GGGAGGAGGG CGTGTTGCGC GAAGGAAGGC 50

GTTAAGAGAG GGGCGTGGAC GAGCAAGGAG GACGATGCCT TGGCCGCCTA 100

CGTCAAGGCC CATGGCGAAG GCAAATGGAG GGAAGTGCCC CAGAAAGCCG 150

gtaaaactag ctagtctttt tatttcattt tgggatcata tatatacccc 200

cgaggcaaga ccggaggacg atcacgtgtg tgggtgcagG TTTGCGTCGG 250

TGCGGCAAGA GCTGCCGGCT GCGGTGGCTG AACTACCTCC GGCCCAACAT 300

CAGGCGCGGC AACATCTCCT ACGACGAGGA GGATCTCATC ATCCGCCTCC 350

ACAGGCTCCT CGGCAACAGg tctgtgcagt ggccagtggt gggctagctt 400

attacacgag ctgacgacga ggcgatcgat cgagcgtctg ctgcgaattc 450

atctgttccg gtgtcggccg tgtgagagtg agctcattca tatgtacatg 500

cgtgttggcg cgcagGTGGT CGCTGATTGC AGGCAGGCTG CCTGGCCGAA 550

CAGACAATGA AATCAAGAAC TACTGGAACA GCACGCTGGG CCGGAGGGCA 600

GGCGCCGGCG CCGGCGCCGG CGGCAGCTGG GTCGTCGTCG CGCCGGACAC 650

CGGCTCGCAC GCCACCCCGG CCGCGACGTC GGGCGCCTGC GAGACCGGCC 700

AGAATAGCGC CGCTCATCGC GCGGACCCCG ACTCAGCCGG GACGACGACG 750

ACCTCGGCGG CGGCGGTGTG GGCGCCCAAG GCCGTGCGGT GCACGGGCGG 800

ACTCTTCTTC TTCCACCGGG ACACGACGCC GGCGCACGCG GGCGAGACGG 850

CGACGCCAAT GGCCGGTGGA GGTGGAGGAG GAGGAGGAGA AGCAGGGTCG 900

TCGGACGACT GCAGCTCGGC GGCGTCGGTA TCGCTTCGCG TCGGAAGCCA 950

CGACGAGCCG TGCTTCTCCG GCGACGGTGA CGGCGACTGG ATGGACGACG 1000

TGAGGGCCCT GGCGTCGTTT CTCGAGTCCG ACGAGGACTG GCTCCGCTGT 1050

CAGACGGCCG GGCAGCTTGC GTAGACAACA AGTACACGTA TAGATGTCCA 1100

ATAAGCACGA GGCCCGCGAG CCCGGCACGA AGCCCGCTTT TTGGGCCCGG 1150

TCCGAGCCCG GCACGGCCCG GTTATATGCA GACCCGGGCC GGCCCGGCAC 1200

GAATAAGCGG GCCGGGCTCG GACAGGAAAT TAGGCACGGT GAGCTAGCCC 1250

GGCACGGCCC GTTTAGGTCT AAGCCCGTTA AGCCCGTTTT TTTACACTAA 1300

AACGTGCTTC TCGGCCCGCA TAGCCCGCTT CTCGGCCCGC TTTTTTCGTG 1350

CTAAACGGGC CGGCCCGGCC CGGTTTAGGC CCGTTGCGGG CCGGGCTCGG 1400

ACAGGAAATT GAGCCCGCGT GCTTAGCCGG CCCGGCCCGG TTTTTTAATC 1450

GTGCCTGGCG GGCCAGGCCC AAAACGGGCC GGGCTTCACC GGGCCCGGGC 1500

CGGACCGGGC CGGGCGGCCC GTTTGGACAT CTCTAAGTAC ACGTATGGAG 1550

GAGAATATAT ATATAGTCAT GCGTACGTAT AGATTTTTTC ATCCGATCCC 1600

AACAGAAATA CGTATGAAAA TGCTCTTCGT TCTTTTTCAT TTATCATATC 1650

TATACTATAC TTAAAACACC AGTTTCAACG GTCGTCATGC GTCATTTTTT 1700

TACAAATAAC CCCTCACAGC TATTTCAAAT TAATCCGCTG CACGTCTATA 1750

GATGCCAAAC GACGCCCAAC ACGGGCTAGA TGCACGCGGG CCACAACTAT 1800

GGCACAGGCA CGTCATGCCG GCCTGCTAAC TGTGTCGGGC TAGCCCGTTA 1850

GCCCGTCGAT CCATTTAATT AAATTAGCGT AACGACGCCC GACACGGGCT 1900

AGATGCACGT GGGCCACAAC TATGGCACAT GCACGTCATG CCGGCCTGTT 1950

AACTGTGTCG GGCCAGTCTG TTAGCCCATT GATCCATTTA ATTAAATCAG 2000

CGTAAAATGT TAAAAACGGT GCAGGAGGTG GGGTTCGAAC CCATACCCTG 2050

ATGGAAGAAG GGCGGGAGAC ACTGGGTGAA ACTGTCTAAC CAGTAGAATA 2100

TCTATCACGC TAAGATGTTT TTAATATTGA ATATAAATTG TATATAAGCA 2150

TATAAGTTTT TTTGTAAAAT AAAAAATAAT CGTGTCGGGC CGGGCCATCA 2200

CTACTGGCCG AGGCTACAAC CCAAGCACGA CACGACGTTC TTGGCTCTTG 2250

CAAGCATTAG GTCGTTTCTG AGACCATATT GGCGCAATGG ACTACATGAT 2300

GTTTGGGGTT GCTGAATTGA ATGGAGCAGC AATAATTTGT CACACTAACA 2350

GCAAAATGAA AGGTTATTTG TTGGTTTTAA ACGTTAGTAA TTGCTACGAA 2400

GTAGCATAAT TTATATGGAG CGCATCCAGT TTTTATTGAT GCCTGACTTT 2450

AGCAATCACT CCATATTTTG ATCTATCTTT TTTATAAGTT TGACTTCATG 2500

GGACTTATTT TAGAACTTGA TCTCACAAAC TTTCTCTTAT TTTGTCTCTA 2550

TATGATGAAA TTGTGTCATT TTATAATCTT TGTTCATTCA GTCAATCGTT 2600

GTGAACTCTC TTCTAATCAC TCACTTCATT AGTTGTGTTG TACCAAGACA 2650

TATTTGCATA GAGTAAACAA TAACATCAGT TAGCCAAATC AAAAAATATA 2700

TTATACAGAG AGCGGAGACA ATCAAATAAA AAATCTTGAA ATTTTTTTAA 2750

TGGATAGTTT ACGTGGGTAT TGTTGTAAGC CGTCGCAACG CACGGGCAAC 2800

CGACTAGTTT TAGTTTATAA ATTAATAAAC GTACGACAAA TATTAAGAAC 2850

GCCACCTTTC CATGCCTACG CGCGCGTGAG ACACGACCGG GGCACGTCAG 2900

CACGTGTGCC CCTGTTGTAT AATTTATTTA CTTTTTAATG ACTATGTGCT 2950

GTTGGTTGCC GTTGGCTTCA TCGTGTTCGT AGCCATGCAT AAATCCAGCG 3000

CCGTACATGT CGATAGAGAA TACTTGCTCT TTTTCAACAA AAAAAGGGGT 3050

AAACTGCACT ATGTATATAG TTATAATAAT AAAATACATA CGATGAGGGA 3100

GTTATTTAAT TTAATTCATC ATTGTAACGT TGTACGCAAT ATTAGAGATT 3150

TATAGATATT TATTATTACT AGATGTGCGT AATACCTTAG CTAGCGCTGA 3200

AGGCTTCTGT AATATCCTTC AACTTGTGTT TACCAAATTC CCCTCTTCAG 3250

TTGAGACAAA TCCACGTGTA GCCAGGAGGG AGTTATTTAA TTTACTCCAC 3300

ATGTGTGAAC ACATGGTTGA TAAGTATCCA TGAGTATCGG TACGAGGAGC 3350

AATTGTCCAC AAAAAAAAAC AAGTTTTGAT TCTGAGGTTA GCCATATCGT 3400

GAGTGTGGTA GCGTACTAGT GTGCAATGTC TTTTAGCTTG ACTATTTATG 3450

ACAATTTAAC CCGTAAACCT GGAACTCGAT GGATGGATAT CTGATCCGAT 3500

GGTTAEGGGT AACGGTGAAG ATTTTGACCT GCGGGTATAG ATGGCCAATA 3550

AGCACGAGGC CCGCGAGCCC GGCACGAAGC CCGCTGTTTG AGCCCGGTCC 3600

GAGCCCGGCA GACCCGATTC TATGCGGGTC CGGGXXGGCC CAGCACGAAT 3650

AAGCGGGTCT GGCTCGGACA GGAAATTAGG CACGACCTGT TTACCTCTAA 3700

GCCCGTTAGG CCCGCTTTTT GCACTAAAAC ATGCTTACCA GCCCGCTTAG 3750

CGCGCTTTTT GGCCCGCTTT TTTCGTGCTA AACGGGCCGG GTCGGCGTCT 3800

TTAGGCCCGC TGCGAGCCGG ACTCGGACAG GAAATCGAGT CCGCGGGCTT 3850

AAACAGCTTG GCCCGATTTT CTAACCGTGC CTGGTGGGTC GGTCCAAAAC 3900

GGGCCGGGCT TCACCGGGCC CGGG 3924

7. BIBLIOGRAPHY

Acheson, N. H. 1984. Kinetics and efficiency of polyadenylation of late polyomavirus nuclear RNA: Generation of oligomeric polyadenylated RNAs and their processing into mRNA. Mol. Cell. Biol. 4:722-729.

Biedenkapp, H., U. Borgmeyer, A. E. Sippel, and K-H. Klempnauer. 1988. Viral myb oncogene encodes a sequence-specific DNA-binding activity. Nature 335:835-837.

Bingham, P. M., R. Levis, G. M. Rubin. 1981. Cloning of DNA sequences from the white locus of D. melanogaster by a novel and general method. Cell 25:693-704.

Boyle, W. J., M. A. Lampert, J. S. Lipsick, and M. A. Baluda. 1984. Avian myeloblastosis and E26 virus oncogene products are nuclear proteins. Proc. Natl. Acad. Sci. USA 81:4265-4269.

Brent, R., and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729-736.

Chappell, J., and K. Hahlbrock. 1984. Transcription of plant defense genes in response to UV light or fungal elicitor. Nature 311:76-78.

Chen, S. 1973. Conditional colored and colorless alleles at the C1 locus. Maize Genet. Coop. Newsl. 47:140-143.

Chen, S., and E. H. Coe. 1977. Control of anthocyanin synthesis by the C locus in maize. Biochem. Genet. 15:333-346.

Choi, O-R., and J. D. Engel. 1986. A 3' enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult β-globin gene. Nature 323:731-734.

Church, G. M., and W. Gilbert. 1984. Genomic sequencing. Proc. Natl. Acad. Sci. 83:1991-1995.

Coe, E. H. 1962. Spontaneous mutation of the aleurone color inhibitor on maize. Genetics 47:779-783.

Coe, E. H. 1964. Compund versus bifunctional nature of the C locus in maize. Genetics 50:571-578.

Coe, E. H. 1982. Genetic interactions of c2 and whp and their duplicate function. Maize Genet. Coop. Newsl. 56:49-50.

Coe, E. H., and M. G. Neuffer. 1977. The genetics of corn. Pp. 111-123. In G. F. Sprague (ed.), Corn and corn improvement. American Society of Agronomy, Inc., Madison, Wisconsin.

Coe, E. H., S. M. McCormick, and S. A. Modena. 1981. White pollen. J. of Hered. 72:318-320.

Cone, K. C., F. A. Burr, and B. Burr. 1986. Molecular analysis of the maize anthocyanin regulatory locus C1. Proc. Natl. Acad. Sci. 83:9631-9635.

Connelly, S., and J. L. Manley. 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes and Development 2:440-452.

Dangl, J. L., K. Hahlbrock, J. Schell. 1989. Structure and regulation of light-inducible genes: Regulation and structure of chalcone synthase genes. In I. K. Vasil, and J. Shell (ed.), Cell culture and somatic cell genetics of plants: Plant Nuclear Genes and their Expression 6:in press.

Dellaporta, S. L., I. Greenblatt, J. L. Kermicle, J. B. Hicks, and S. R. Wessler. 1988. Pp. 263-282. In J. P. Gustafson and R. Appels (eds.), Chromosome structure and function: Impact of new concepts. Plenum Press, New York.

Dooner, H. K. 1983. Coordinate genetic regulation of flavonoid biosynthesic enzymes in maize. Mol. Gen. Genet. 189:136-141.

Dooner, H. K. 1985. Vivipaous-1 mutation in maize conditions pleiotropic enzyme deficiencies in the aleurone. Plant Physiol. 77:486-488.

Dooner, H. K., and O. E. Nelson. 1977. Genetic control of UDPglucose: Flavonol 3-O-glucosyltransferase in the endosperm of maize. Biochem. Genet. 15:509-519.

Dooner, H. K., and O. E. Nelson. 1979. Interaction among C, R and Vp in the control of the Bz glucosyltransferase during endosperm development in maize. Genetics 91:309-315.

Döring, H. P., and P. Starlinger. 1984. Barbara McClintock's controlling elements: Now at the DNA level. Cell 39:253-259.

Döring, H. P., E. Tillmann, and P. Starlinger. 1984. DNA sequence of the maize transposable element Dissociation. Nature 307:127-130.

Fedoroff, N., S. Wessler, and M. Shure. 1983. Isolation of the transposable maize elements Ac and Ds. Cell 35:235-242.

Fincham, J. R. S., and G. R. K. Sastry. 1974. Controlling elements in maize. Ann. Rev. Genet. 8:15-50.

Fischer, J. A., and T. Maniatis. 1986. Regulatory elements involved in Drosophila Adh gene expression are conserved in divergent species and seperate elements mediate expression in different tissues. EMBO J. 5:1275-1289.

Förger, D., 1988. Isolierung und Charakterisierung von Mutanten des c1-Locus aus Zea mays. Unpublished Diplom Thesis. Universität Köln, Köln, FRG.

Gil, A., and N. J. Proudfoot. 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA 3'end formation. Cell 49:399-406.

Gill, G., and M. Ptashne. 1987. Mutants of GAL4 protein altered in an activation function. Cell 51:121-126

Gill, G., and M. Ptashne. 1988. Negative effect of the transcriptional activator GAL4. Nature 334:721-724.

Giniger, E., and M. Ptashne. 1987. Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. Nature 330:670-672.

Hagenbüchle, O., P. K. Wellauer, D. L. Cribbs, and U. Schibler. 1984. Termination of transcription in the mouse α-amylase gene amy-2a occurs at multiple sites downstream of the polyadenylation site. Cell 38:737-744.

Heidecker, G., and J. Messing. 1986. Structural analysis of plant genes. Ann. Rev. Plant. Physiol. 37:439-466.

Henikoff, S. 1987. Unidirectional digestion with the exonuclease III in DNA sequence analysis. Methods in Enzymology 155:156-165.

Henikoff, S., and F. Hutchinson. 1987. Exonuclease III generated deletions for DNA sequence analysis. Promega Notes 8:1-3.

Higgs, D. R., S. E. Y. Goodbourn, J. Lamb, J. B. Clegg, D. J. Weatherall, and N. J. Proudfoot. 1983. α-Thalassaemia caused by a polyadenylation signal mutation. Nature 306:398-400.

Hope, I. A., S. Mahadevan, and K. Struhl. 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature 333:635-640.

Hsu, S. M. 1970. The development of pigments in germinating colorless seeds. Maize Genet. Coop. Newsl. 44:153-154.

Johnson, M. R., C. Norman, M. A. Reeve, J. Scully, and N. J. Proudfoot. 1986. Tripartite sequences within and 3' to the sea urchin H2A histone gene display properties associated with a transcriptional termination process. Mol. Cell. Biol. 6:4008-4018.

Johnston, S. A., J. M. Salmeron Jr., and S. S. Dincher. 1987. Interaction of positive and negative regulatory proteins in the galactose regulon of yeast. Cell 50:143-146.

Kiesselbach, T. A. 1980. The structure and reproduction of corn. University of Nebraska Press, Lincoln, Nebraska.

Kirby, L. T., and E. D. Styles. 1970. Flavonoids associated with specific gene action in maize aleurone, and the role of light in substituting for the action of a gene. Can. J. Genet. Cytol. 12:934-940.

Klempnauer, K-H., and A. E. Sippel. 1987. The highly conserved amino-terminal region of the protein encoded by the v-myb oncogene functions as a DNA-binding domain. EMBO J. 6:2719-2725.

Kreuzaler, F., H. Ragg, E. Fautz, D. Kuhn, and K. Hahlbrock. 1983. UV-Induction of chalcone synthase mRNA in cell suspension cultures of Petroselinum hortense. Proc. Natl. Acad. Sci. U.S.A. 80:2591-2593.

Lanoix, J., and N. H. Acheson. 1988. A rabbit β-globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA. EMBO J. 7:2515-2522.

Levitt, N., D. Briggs, A. Gil, and N. J. Proudfoot. 1989. Definition of an efficient synthetic poly(A) site. Genes and Development 3:1019-1025.

Lin, Y. S., M. F. Carey, M. Ptashne, and M. R. Green. 1988. GAL4 derivatives function alone and synergistically with mammalian activators in vitro. Cell 54:659-664.

Lue, N. F., D. I. Chasman, A. R. Buchman, and R. D. Kornberg. 1987. Interaction of GAL4 and GAL80 gene regulatory proteins in vitro. Mol. Cell. Biol. 7:3446-3451.

Ma, J., and M. Ptashne. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell 48:847-853.

Maniatis, T., E. F. Fritisch, and J. Sambrook. 1984. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Maxam, A. M., and W. Gilbert. 1980. Seqencing end-labeled DNA with base specific chemical cleavages. Methods in Enzymology 65:499-560.

McCarty, D. R., C. B. Carson, P.S. Stinard, and D. S. Robertson. 1989. Molecular analysis of viviparous-1: An abscisic acid insensitive mutant of maize. The Plant Cell 1:523-532.

McClintock, B. 1946. Maize genetics. Carnegie Inst. Wash. Year Book 45:176-186.

McClintock, B. 1947. Cytogenetic studies of maize and neurospora. Carnegie Inst. Wash. Year Book 46:146-152.

McClintock, B. 1948. Mutable loci in maize. Carnegie Inst. Wash. Year Book 47:155-169.

McClintock, B. 1949. Mutable loci in maize. Carnegie Inst. Wash. Year Book 48:142-154.

McClintock, B. 1950. Mutable loci in maize. Carnegie Inst. Wash. Year Book 49:157-167.

McClintock, B. 1951. Chromosome organization and genic expression. Cold Spring Harbor Symp. Quan. Biol. 16:13-47.

McCormick, S. 1977a. C-I and light. Maize Genet. Coop. Newsl. 53:31-32.

McCormick, S. 1977b. Pigment development on immature c-p ears. Maize Genet. Coop. Newsl. 53:32-33.

McCormick, S. 1977c. Light inducibility of anthocyanins in vp aleurone. Maize Genet. Coop. Newsl. 53:33-35.

McCormick, S. 1978. Pigment synthesis in maize aleurone from precursors fed to anthocyanin mutants. Biochem. Genet. 16:777-785.

Neuffer, M. G., L. Jones, and M. S. Zuber. 1968. The mutants of maize. Crop Science Society of America, Madison, Wisconsin.

Nevers, P., N. S. Shepherd, and H. Saedler. 1986. Plant transposable elements. Advances in Botanical Research 12:103-203.

Niesbach-Klösgen, U. 1987. Molekulare Analyse des C2 Gens aus Zea mays L. und Studien zur Evolution der Chalkonsynthase in Pflanzen. Unpublished Ph.D. Thesis. Universität Köln, Köln, FRG.

Niesbach-Klösgen, U., E. Barzen, J. Bernhardt, W. Rohde, Zs. Schwarz-Sommer, H. J. Reif, U. Wienand, and H. Saedler. 1987. Chalcone synthase genes in plants: a tool to study evolutionary relationships. J. Mol. Evol. 26:213-225.

O'Reilly, C., N. S. Shepherd, A. Pereira, Z. Schwarz-Sommer, I. Bertram, D. S. Robertson, P. A. Peterson, and H. Saedler. 1985. Molecular cloning of the a1 locus of Zea mays using the transposable elements En and Mu1. EMBO J. 4:877-882.

Paz-Ares, J., U. Wienand, P. A. Peterson, and H. Saedler. 1986. Molecular cloning of the c locus of Zea mays: A locus regulating the anthocyanin pathway. EMBO J. 5:829-833.

Paz-Ares, J., D. Ghosal, U. Wienand, P. A. Peterson, and H. Saedler. 1987. The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. 6:3553-3558.

Paz-Ares, J., D. Ghosal, and H. Saedler. 1989. Molecular analysis of the C1-I allele from Zea mays: A dominant mutant of the regulatory C1 locus. In preparation.

Pereira, A., Z. Schwarz-Sommer, A. Gierl, I. Bertram, P. A. Peterson, and H. Saedler. 1985. Genetic and molecular analysis of the Enhancer (En) transposable element system of Zea mays. EMBO J. 4:17-23.

Peterson, P. A. 1978. Controlling elements: The induction of mutability at the A2 and C loci in maize. Pp. 601-631. In D. B. Walden (ed.), Maize Breeding and Genetics. John Wiley and Sons, Inc., New York.

Peterson, P. A. 1988. The evolutionary role of mobile element transpositions in plants. Pp. 31-47. In D. E. Styles, G. A. Gavazzi, and M. L. Racchi (eds.), Proceedings of a Post Congress Meeting of the XVI International Congress of Genetics.

Pribyl, T. M., and H. G. Martinson. 1988. Transcription termination at the chicken β^H-glogin gene. Mol. Cell. Biol. 8:5369-5377.

Proudfoot, N. J. 1989. How RNA ploymerase II terminates transcription in higher eukaryotes. TIBS 14:105-110.

Proudfoot, N. J., and E. Whitelaw. 1988. Frontiers in Molecular Biology. Pp. 97-129. In B. D. Hames and D. M. Glover (eds.), Molecular Immunology. IRL Press, Washington, D.C.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. Nature 335:683-689.

Putney, S. D., S. J. Benkovic, and P. R. Schimmel. 1981. A DNA fragment with an α-phophorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. Proc. Natl. Acad. Sci. U.S.A. 78:7350-7354.

Reddy, G. M., and E. H. Coe. 1962. Inter-tissue complementation: A simple technique for direct anaylsis of gene-action sequence. Science 138:149-150.

Sadowski, I., J. Ma, and M. Ptashne. 1988. Gal4-VP16 is an unusually potent transcriptional activator. Nature 335:563-564.

Saedler, H., and P. Nevers. 1985. Transposition in plants: a molecular model. EMBO J. 4:585-590.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. (USA) 74:5463-5467.

Scheffler, B. E. 1986. Genetic characterization of vp-m451: A mutable viviparous allele controlled by the Cy transposable element in Zea mays L. Unpublished M.S. Thesis. Iowa State University, Ames, Iowa.

Schröder, J., F. Kreuzaler, E. Schäfer, and K. Hahlbrock. 1979. Concomitant induction of phenylalanine ammonia-lyase and flavone synthase mRNAs in irradiated plant cells. J. Biol. Chem. 254:57-65.

Schulze-Lefert, P., J. L. Dangl, M. Becker-Andre, K. Hahlbrock, and W. Schulz. 1989. Inducible *in vivo* DNA footprints define sequences necessary for UV light activation of the parsley chalcone synthase gene. EMBO J. 8:651-656.

Schwarz-Sommer, Zs., A. Gierl, R. B. Klösgen, U. Wienand, P. A. Peterson, and H. Saedler. 1984. The *Spm* (*En*) transposable element controls the excision of a 2 kb DNA insert at the *wx-m8* locus of *Zea mays*. EMBO J. 3:1021-1028.

Schwarz-Sommer, Z., A. Gierl, R. B. Berndtgen, and H. Saedler. 1985. Sequence comparison of "states" of *a1-m1* suggest a model of Spm (En) action. EMBO J. 4:2439-2443.

Struhl, K. 1987. Promoters, activator proteins and the mechanism of transcriptional initiation in yeast. Cell 49:295-297.

Styles, E. D., and O. Ceska. 1977. The genetic control of flavonoid synthesis in maize. Can. J. Genet. Cytol. 19:289-302.

Sutton, W. D., W. L. Gerlach, D. Schwarz, and W. J. Peacock. 1984. Molecular analysis of *Ds* controlling element mutations at the *Adh1* locus of maize. Science 233:1265-1268.

Trainor, C. D., S. J. Stamler, and J. D. Engel. 1987. Erthroid-specific transcription of the chicken histone H5 gene is directed by a 3' enhancer. Nature 328:827-830.

Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259-268.

Walbot, V., and J. Messing. 1988. Molecular genetics of corn. Pp. 111-123. *In* G. F. Sprague and J. W. Dudley (eds.), Corn and corn improvement. American Society of Agronomy, Inc., Madison, Wisconsin.

Weck, E., U. Courage, H. P. Döring, N. Federoff, and P. Starlinger. Analysis of sh-m6233, a mutation induced by the transposable element Ds in the sucrose gene of Zea mays. EMBO J. 3:1713-1716.

Wellmann, E. 1975. Eine quantitative Analyse des Lichteffekts auf die Flavonoidsynthese in pflanzlichen Zell- und Gewebekulturen. Planta Medica, Suppl. 107-111.

Whitelaw, E., and N. Proudfoot. 1986. α-Thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human α2 globin gene. EMBO J. 5:2915-2922.

Wienand, U., and H. Saedler. 1987. Plant transposable elements: Unique structures for gene tagging and gene cloning. Pp. 205-227. In Th. Hohn and J. Schell (eds), Plant DNA infectious agents. Springer-Verlag Wien, New York.

Wienand, U., J. Paz-Ares, B. E. Scheffler, and H. Saedler. 1989. Molecular analysis of gene regulation in the anthocyanin pathway in Zea mays. In C. Lamb and R. Beachy (eds.), Plant Gene Transfer: UCLA symposi on Molecular and Cellular Biology. Alan R. Liss, Inc., New York. In press.

8. ACKNOWLEDGMENTS

I would like to thank Dr. Peter A. Peterson for all the support and educational opportunities he has provided me throughout my graduate program. During the time I spent working in his lab I was able to fulfill one of my childhood dreams, and that was to become a geneticist. Although I learned how to do genetics in his lab, the most important training I received from him was how to identify scientific questions and how to design experiments in order to answer them. In other words, he taught me how to think properly, which is something many graduate students never have the opportunity to learn.

I am grateful to Dr. Heinz Saedler for accepting me into his department at the Max-Planck-Institut für Züchtungsforschung, Köln, Federal Republic of Germany, to do my doctoral research. Not only did he obtain funding for me through the Max-Planck-Society, but he also supported my work at the institute with helpful discussions.

My sincerest thanks go to Dr. Udo Wienand who was my project leader at the Max-Planck-Institut. Udo is without a doubt one of the most patient and kind persons I have ever meet. He took the time to teach me the basics of molecular biology and some of the general principles of

soccer. I feel most fortunate that I can call him my friend.

Over the years I have had the privilege to work with several people in Köln and in Dr. Peterson's lab that have given me both moral and technical support. Unfortunately, the number of people is so large that it prevents me from mentioning all of them. Therefore, I will only say that I am indebted to everyone of them. Of all those people, I give special thanks to Dr. Javier Paz-Ares who was more than generous with his help.

I wish to thank the members of my advisory committee, Dr. Arnel Hallauer, Dr. Walter Fehr, Dr. Stephen Gendel, Dr. John Mayfield, and Dr. David Morris.

To my parents, who have given me their constant support throughout my education, I give my most sincerest thanks.

My greatest thanks goes to my wife Jodi, who stood by me through the good and bad times even before we were married.